Limit theorems for orthogonal polynomials related to circular ensembles

Najnudel, Joseph ; Nikeghbali, Ashkan ; Rouault, Alain

Abstract: For a natural extension of the circular unitary ensemble of order n, we study as $n \to \infty$ the asymptotic behavior of the sequence of monic orthogonal polynomials $(\Phi_{k,n},\ k = 0, \dots ,n)$ with respect to the spectral measure associated with a fixed vector, the last term being the characteristic polynomial. We show that, $n \to \infty$, the sequence of processes $(\log \Phi_{\lfloor nt \rfloor,n}\ (1),\ t \in [0,1])$ converges to a deterministic limit, and we describe the fluctuations and the large deviations.

DOI: https://doi.org/10.1007/s10959-015-0632-x

Originally published at:
Najnudel, Joseph; Nikeghbali, Ashkan; Rouault, Alain (2016). Limit theorems for orthogonal polynomials related to circular ensembles. Journal of Theoretical Probability, 29(4):1199-1239.
DOI: https://doi.org/10.1007/s10959-015-0632-x

LIMIT THEOREMS FOR ORTHOGONAL POLYNOMIALS RELATED TO CIRCULAR ENSEMBLES

J. NAJNUDEL, A. NIKEGHBALI, AND A. ROUAULT

ABSTRACT. For a natural extension of the circular unitary ensemble of order n, we study as $n \to \infty$ the asymptotic behavior of the sequence of orthogonal polynomials with respect to the spectral measure. The last term of this sequence is the characteristic polynomial. After taking logarithm and rescaling, we obtain a process indexed by $t \in [0, 1]$. We show that it converges to a deterministic limit, and we describe the fluctuations and the large deviations.

1. INTRODUCTION

In this paper, we study the asymptotic behavior of a triangular array of complex random variables originating from Random Matrix Theory, and more specifically from some models of unitary ensembles. Our results can be viewed as an extension of some earlier works on the characteristic polynomial of unitary matrices. Indeed, for fixed n, our random variables consist of the sequence of monic orthogonal polynomials with respect to a random measure associated with a random unitary matrix of size n, when evaluated at the point $z = 1$, the last term being the characteristic polynomial of that matrix.

It was already noticed in [3] that the characteristic polynomial of a random unitary matrix sampled from the Haar measure on $\mathbb{U}(n)$, the unitary group of order n has the same law as a product of independent random variables. In a previous paper ([5]) we saw that this characteristic polynomial $\Phi_n(z)$ (evaluated at $z = 1$) is actually a product of variables

$$\Phi_n(1) = \prod_{j=0}^{n-1}(1 - \gamma_j)$$

and we named the independent variables γ_j the modified Verblunsky coefficients, referring to the coefficients involved in the Schur recursion for orthogonal polynomials on the unit circle (OPUC), as dubbed by Simon [25]. These variables (for $j < n$) have an explicit density in the open unit disc, depending on their rank j. We proved in [5] that this description stays valid when we change the probability measure on $\mathbb{U}(n)$, by considering the Circular Jacobi Ensemble, that we will define below. The construction of the deformed Verblunsky coefficients uses the whole sequence $\{\Phi_{k,n}(z), k = 0, \dots, n\}$ of monic orthogonal polynomials with respect to the spectral measure of the pair (U, e_1) where e_1 is a fixed vector. We have ([5]

2010 *Mathematics Subject Classification.* 15B52, 42C05, 60F10, 60F17 .

Key words and phrases. Random Matrices, unitary ensemble, orthogonal polynomials, large deviation principle, invariance principle.

Prop. 2.2)

$$\Phi_{k,n}(1) = \prod_{j=0}^{k-1}(1-\gamma_j)$$

It is then natural to study the triangular array $\{\Phi_{k,n}(1), k = 0, \dots, n\}$ and in order to normalize the time for different values of n, one can study the *process* $\{\Phi_{\lfloor nt \rfloor,n}(1), t \in [0,1]\}$, where $\lfloor \ \rfloor$ denotes the integer part, and where, by convention, $\Phi_{0,n}(1) = 1$.

One can note that $\Phi_{n,n}$ is the characteristic polynomial of U and that when U is the chosen according to the Haar measure, the sequence of random variables $(\Phi_{n,n}(1))$ has played a crucial role in the recent interactions between random matrix theory and number theory (see [14]). It is also worthwhile to note that in [17], Killip and Stoiciu considered also a stochastic process indexed by $t = k/n$ and related to the sequence of orthogonal polynomials in the CβE ensemble. In fact they considered variables which are the complex conjugate of our deformed Verblunsky coefficients as auxiliary variables in the study of the Prüfer phase (Lemma 2.1 in [17]). This tool is used again in [15]. Later Ryckman ([23] Section 4) used a version of the deformed Verblunsky coefficients in the proof of its joint asymptotic laws.

To be more precise and to explain the interest of our approach, we now describe our model. For $n \geq 1$, $\beta > 0$ and $\delta \in \mathbb{C}$ such that $\mathfrak{Re}\, \delta > -1/2$, we consider a distribution $\mathrm{CJ}^{(n)}_{\beta,\delta}$ on the set of probability measures on the unit circle $\mathbb{T}$, supported by n points. This family of distribution generalizes the Circular Jacobi Ensemble (the notation CJ comes from this fact), and it can be defined as follows. If the random measure

$$\mu = \sum_{j=1}^{n} \pi_j \, \delta_{e^{i\theta_j}}\,,$$

for $\theta_j \in [0, 2\pi)$, has the distribution $\mathrm{CJ}^{(n)}_{\beta,\delta}$, then:

- The joint density $h^{(n)}_{\delta,\beta}$ of the law of $(\theta_1, \dots, \theta_n)$, with respect to the Lebesgue measure on $[0, 2\pi)^n$, is given by

$$h^{(n)}_{\delta,\beta}(\theta_1, \dots, \theta_n) = c^{(n)}_{\delta,\beta} |\Delta(e^{\mathrm{i}\theta_1}, \dots, e^{\mathrm{i}\theta_n})|^{\beta} \prod_{j=1}^{n} (1 - e^{-\mathrm{i}\theta_j})^{\delta} (1 - e^{\mathrm{i}\theta_j})^{\overline{\delta}}, \tag{1.1}$$

  where $c^{(n)}_{\delta,\beta} > 0$ is a normalization constant.
- The weights $(\pi_1, \dots, \pi_n)$ follow a Dirichlet law of parameter β' on the simplex $\pi_1 + \cdots + \pi_n = 1$, $\pi_j > 0$, where $\beta' := \beta/2$ (we conserve this notation β' in all the sequel of the paper).
- The tuples $(\theta_1, \dots, \theta_n)$ and $(\pi_1, \dots, \pi_n)$ are independent.

Besides, when U is an unitary matrix and e_1 a cyclic vector for U, then the spectral measure μ of the pair (U, e_1) is defined as the unique (probability) measure on $\mathbb{T}$ such that

$$\langle e_1, U^j e_1 \rangle = \int_{\mathbb{T}} z^j \mu(dz) \quad j \in \mathbb{Z}\,.$$

When $\delta = 0$, it was proved in [16] that the distribution of the spectral measure of the pair (U, e_1) where U is randomly sampled from $\mathbb{U}(n)$ according to the Haar measure and e_1 is a fixed vector of $\mathbb{C}^n$, for instance $(1, 0, \dots, 0)$, is precisely $\mathrm{CJ}^{(n)}_{2,0}$. For $\delta = 0$ and $\beta > 0$, these authors found a model of random unitary matrices such that the spectral measure has the distribution $\mathrm{CJ}_{\beta,0}^{(n)}$. In [5], we gave a model corresponding to the case $\delta \neq 0$. All these constructions rely on the theory of orthogonal polynomials on the unit circle (OPUC) that we recall now.

From the linearly independent family of monomials $\{1, z, z^2, \dots, z^{n-1}\}$ in $L^2(\mathbb{T}, \mu)$, we construct an orthogonal basis $\Phi_{0,n}, \dots, \Phi_{n-1,n}$ of monic polynomials by the Gram-Schmidt procedure. The n^{th} degree polynomial obtained this way is

$$\Phi_n(z) = \Phi_{n,n}(z) = \prod_{j=1}^{n} (z - e^{i\theta_j}),$$

i.e. the characteristic polynomial of U. The Φ_k's ($k = 0, \dots, n$) obey the Szegö (or Schur) recursion relation:

$$\Phi_{j+1,n}(z) = z\Phi_{j,n}(z) - \bar{\alpha}_j \Phi_{j,n}^*(z)$$

where

$$\Phi_{j,n}^*(z) = z^j \, \overline{\Phi_{j,n}(\bar{z}^{-1})}\,.$$

The coefficients α_j ($j \geq 0$) are called Schur or Verblunsky coefficients and satisfy the condition $\alpha_0, \cdots, \alpha_{n-2} \in \mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ and $\alpha_{n-1} \in \mathbb{T}$. There is a bijection between this set of coefficients and the set of spectral probability measures ν (Verblunsky's theorem). We can write the orthogonal polynomials with the help of a new system of functions built from the Verblunsky coefficients. Setting

$$y_k(z) = z - \frac{\Phi_{k+1,n}(z)}{\Phi_{k,n}(z)} = \bar{\alpha}_k \frac{\Phi_{k,n}^*(z)}{\Phi_{k,n}(z)}\ , \quad (k = 0, \dots, n-1)\,,$$

we have $y_0(z) = \bar{\alpha}_0$ and the following decomposition:

$$\Phi_{k,n}(z) = \prod_{j=0}^{k-1} (z - y_j(z))\ , \quad k = 1, \dots, n\,.$$

If $\gamma_j := y_j(1)$, we get

$$\Phi_{k,n}(1) = \prod_{j=0}^{k-1} (1 - \gamma_j)\ ,\ k = 1, \dots n \tag{1.2}$$

and in particular

$$\det(I - U) = \Phi_n(1) = \Phi_{n,n}(1) = \prod_{j=0}^{n-1} (1 - \gamma_j)\,. \tag{1.3}$$

Note that the definition implies $|\gamma_k| = |\alpha_k|$, and in particular $|\gamma_{n-1}| = 1$. In the sequel, following [5], we refer to the γ_j's as the *deformed Verblunsky coefficients*.

In [5], it is proven that for μ following the distribution $\mathrm{CJ}_{\beta,\delta}^{(n)}$, the coefficients $(\gamma_j)_{0 \leq j \leq n-1}$ are independent (note that in general it is not true for the classical Verblunsky coefficients, except if $\delta = 0$) and their distributions are explicitly computable. More precisely, for $r > 0$ and $\delta \in \mathbb{C}$ such that $r + 2\mathfrak{Re}\,\delta + 1 > 0$, let $g_r^{(\delta)}$ be the density on the unit disc $\mathbb{D}$ proportional to

$$\left(1 - |z|^2\right)^{r-1} (1 - z)^{\bar{\delta}} (1 - \bar{z})^{\delta}$$

and let $\lambda^{(\delta)}$ be the density on the unit circle $\mathbb{U}$ proportional to

$$(1 - z)^{\bar{\delta}} (1 - \bar{z})^{\delta}\,.$$

Then, for $j < n-1$, γ_j has density $g^{(\delta)}_{\beta'(n-j-1)}$ and γ_{n-1} has density $\lambda^{(\delta)}$. Note that the law of γ_{n-1} does not depend on n and β. The values of the normalization constants can be easily deduced from the computation of integrals which are collected in the appendix of the present paper.

Let us now explain why we choose to focus on the value at 1 of the characteristic polynomial (in addition to its major role in the number theory connections mentioned earlier). Note that in the case $\delta = 0$ (which corresponds to the classical Circular Jacobi Ensemble, and in particular the Circular Unitary Ensemble for $\beta = 2$), the law of $\Phi_n(z)$ does not depend on $z \in \mathbb{T}$, since the distribution of the eigenvalues of U is invariant by rotation. On the other hand, for $\delta \neq 0$, the point 1 plays an important role since it is a singularity of the potential. It is then classical to study the behavior of $\Phi_n(1)$ in the large n asymptotics (see for example [14], [12], [15], [23]). Notice that all these authors consider the case $\delta = 0$.

Our extension to a study of the array $\{\Phi_{k,n}(1), k \leq n\}$ has its own interest as a study of characteristic polynomials. It comes from the following remark. From a measure μ carried by n points, one can also define a $n \times n$ unitary matrix U_n, called GGT by Simon [26] section 10, and which is the matrix of the linear application h on $L^2(\mathbb{T}, \mu)$ given by $h(f)(z) = zf(z)$, taken in a basis of orthonormal polynomials with respect to μ. If $1 \leq k \leq n$, one can denote by $G_k(U_n)$ the $k \times k$ topleft submatrix of U (which is not unitary in general). Then it is known (see for example Forrester [10] Prop. 2.8.2, or Simon [26] proof of Prop. 3.1) that one has for all $k \in \{1, 2, \ldots, n\}$:

$$\Phi_{k,n}(1) = \det(I_k - G_k(U_n)) \tag{1.4}$$

For other aspects of this model see [6].

Most of the results we will obtain on the process $\{\Phi_{\lfloor nt \rfloor, n}(1), t \in [0,1]\}$ will in fact concern its logarithm. Note that even when $\Phi_{k,n}(1) \neq 0$, its complex logarithm is not obvious to define rigorously, since its imaginary part is a priori given only modulo 2π. However, there is a natural way to deal with this issue, which is described in the appendix. One can then fully justify the following formula:

$$\log \Phi_{k,n}(1) = \sum_{j=0}^{k-1} \log(1 - \gamma_j), \tag{1.5}$$

when $\Phi_{k,n}(1) \neq 1$, which occurs almost surely under $\mathrm{CJ}^{(n)}_{\beta,\delta}$.

We study asymptotic properties of this determinant as $n \to \infty$ under essentially two regimes:

• First regime: δ is fixed and $\mathfrak{Re}\,\delta > -1/2$ (hence, this regime includes the case $\delta = 0$).

• Second regime: $\delta = \beta' \mathrm{d} n$ with $\mathfrak{Re}\,\mathrm{d} > 0$.

Some of the results proved here were already announced in [4]. When $\beta = 1, 2, 4$, the independence of the random variables $\Phi_{k+1,n}/\Phi_{k,n}$ $k = 0, \ldots, n-1$ and the identification of their distributions are strongly related to the results of Neretin ([18] Corollary 2.1). In that framework, they can be carried easily on models of matrix balls via his Proposition 2.3. Actually our results may be extended to all models where a remarkable separation of variables occur (see [19]). More precisely the paper is organized as follows.

In Section 2 we study the variables $\log \Phi_{\lfloor nt \rfloor, n}(1)$ for both regimes and prove that their expectations converge to some explicit deterministic function of t, and then we study the fluctuations of $(\log \Phi_{\lfloor nt \rfloor,n}(1))$ as a stochastic process on the space of càdlàg $\mathbb{R}^2$-valued functions. It appears that in the first regime, one has to distinguish between the case $0 < t < 1$ and $t = 1$; at $t = 1$ some transition is occuring since the normalization is changed. Indeed, anticipating the notation of next section where we write $\log \Phi_{\lfloor nt \rfloor,n}(1) - \mathbb{E} \log \Phi_{\lfloor nt \rfloor,n}(1) = \xi_n(t) + \mathrm{i}\eta_n(t)$ and $\zeta_n(t) = \begin{pmatrix} \xi_n(t) \\ \eta_n(t) \end{pmatrix}$, we shall prove that in the first regime, $\{\zeta_n(t);\ t < 1\}$ converges to some explicit Gaussian diffusion $\{\zeta_{\mathrm{d}}^0(t);\ t < 1\}$ while $\zeta_n(1)/\sqrt{n}$ converges to a Gaussian random variable which is independent of the diffusion $\{\zeta_{\mathrm{d}}^0(t);\ t < 1\}$. In the second regime, $\{\zeta_n(t);\ 0 \le t \le 1\}$ converges to some explicit Gaussian diffusion and there is no normalization to perform.

Section 3 is devoted to the establishing Large Deviation Principle (LDP) for the distributions of the real and imaginary parts of $\log \Phi_{\lfloor nt \rfloor,n}(1)$ as a two dimensional random vector with values in the Skorokhod space endowed with the weak topology. We focus our study in the first regime on the case $\delta = 0$. Again in the first regime the case $t = 1$ is playing a special role and is not included. From the contraction principle we are then able to deduce LDP for the marginals at fixed time. Our approach is standard: we first compute the normalized cumulant generating function, compute its limit as well as its dual transform, and end with exponential tightness.

In Section 4 we discuss the connections between the results of Section 3 and existing results on the LDP for the empirical spectral distribution for the circular Jacobi ensemble.

Section 5 gathers in appendix some properties of remarkable functions and densities used in the proofs, in particular the Gamma function Γ and the Digamma function Ψ. For $z \in \mathbb{C}\backslash\mathbb{R}_-$, we take the notation:

$$\ell(z) := \log \Gamma(z)\ ;\ \Psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}\,.$$

The determination of the logarithm in ℓ is chosen in the unique way such that ℓ is holomorphic on $\mathbb{C}\backslash\mathbb{R}_-$ and real on $\mathbb{R}_+^*$.

All along the paper, we use the entropy function $\mathcal{J}$ defined by :

$$\mathcal{J}(u) = \begin{cases} u \log u - u + 1 & \text{if } u > 0 \\ 1 & \text{if } u = 0 \\ +\infty & \text{if } u < 0 \end{cases} \tag{1.6}$$

and its primitive

$$F(t) = \int_0^t \mathcal{J}(u)\, du = \frac{t^2}{2} \log t - \frac{3t^2}{4} + t, \quad (t \ge 0)\,. \tag{1.7}$$

When the arguments of $\mathcal{J}$ or of F are complex, we choose the principal determination of the logarithm.

2. Convergence and fluctuations

Let, for $\mathfrak{Re}\,\mathrm{d} > 0$ and $0 < t \le 1$,

$$\begin{aligned} \mathcal{F}_{\mathrm{d}}(t) &= \log(1 + 2\mathfrak{Re}\,\mathrm{d}) - \log(1 + \bar{\mathrm{d}}) - \log(1 - t + 2\mathfrak{Re}\,\mathrm{d}) + \log(1 - t + \bar{\mathrm{d}}) \\ \mathcal{E}_{\mathrm{d}}(t) &= \mathcal{J}(1 + 2\mathfrak{Re}\,\mathrm{d}) - \mathcal{J}(1 + 2\mathfrak{Re}\,\mathrm{d} - t) - \mathcal{J}(1 + \bar{\mathrm{d}}) + \mathcal{J}(1 + \bar{\mathrm{d}} - t)\,. \end{aligned} \tag{2.1}$$

In this section, we are interested in the process

$$\zeta_n(t) = \begin{pmatrix} \xi_n(t) \\ \eta_n(t) \end{pmatrix}$$

where we have written

$$\log \Phi_{\lfloor nt \rfloor, n}(1) - \mathbb{E} \log \Phi_{\lfloor nt \rfloor, n}(1) = \xi_n(t) + \mathrm{i}\eta_n(t)\,.$$

As a consequence of the result just below, $\mathbb{E} \log \Phi_{\lfloor nt \rfloor, n}(1)$ is finite, and then $\zeta_n(t)$ is well-defined.

2.1. **Convergence to a deterministic limit.**

Theorem 2.1. (1) *In the first regime, i.e. for fixed δ, and for n going to infinity,*

$$\mathbb{E} \log \Phi_n(1) = \frac{\delta}{\beta'} \log n + C + o(1) \tag{2.2}$$

where C is a constant, and for $0 < t < 1$,

$$\mathbb{E} \log \Phi_{\lfloor nt \rfloor, n}(1) \quad = \quad -\frac{\delta}{\beta'} \log(1-t) + o(1)\,. \tag{2.3}$$

(2) *In the second regime, i.e. $\delta = \beta' \mathrm{d} n$, and for $0 < t \leq 1$, we have*

$$\lim_{n\to\infty} \left(\mathbb{E} \log \Phi_{\lfloor nt \rfloor, n}(1) - n \mathcal{E}_{\mathrm{d}} \left(\frac{\lfloor nt \rfloor}{n} \right) \right) = \left(\frac{1}{2} - \frac{1}{\beta} \right) \mathcal{F}_{\mathrm{d}}(t)\,, \tag{2.4}$$

uniformly in t.

(3) *In the first regime, we have*

$$\lim_{n\to\infty} \frac{1}{\log n} \ \mathrm{Cov}\ \zeta_n(1) = \frac{1}{\beta} I_2, \tag{2.5}$$

where I_2 denotes the 2×2 identity matrix, and for $0 < t < 1$,

$$\lim_{n\to\infty} \ \mathrm{Cov}\ \zeta_n(t) = \int_0^t \mathcal{Z}_s^0 ds, \tag{2.6}$$

where

$$\mathcal{Z}_t^0 := \frac{1}{\beta(1-t)} I_2.$$

(4) *In the second regime, we have for $0 < t \leq 1$:*

$$\lim_{n\to\infty} \ \mathrm{Cov}\ \zeta_n(t) = \int_0^t \mathcal{Z}_s^{\mathrm{d}} ds \tag{2.7}$$

where

$$\mathcal{Z}_t^{\mathrm{d}} = \frac{1}{\beta'} \begin{pmatrix} \dfrac{1}{1-t+2\mathfrak{Re}\,\mathrm{d}} - \mathfrak{Re}\, \dfrac{1}{2(1-t+\mathrm{d})} & \mathfrak{Im}\, \dfrac{1}{2(1-t+\mathrm{d})} \\ \mathfrak{Im}\, \dfrac{1}{2(1-t+\mathrm{d})} & \mathfrak{Re}\, \dfrac{1}{2(1-t+\mathrm{d})} \end{pmatrix}. \tag{2.8}$$

(5) *In the second regime,*

$$\sup_{t\in[0,1]} \left| \frac{1}{n} \log \Phi_{\lfloor nt \rfloor, n}(1) - \mathcal{E}_{\mathrm{d}}(t) \right| \underset{n\to\infty}{\longrightarrow} 0 \tag{2.9}$$

in L^2, and then in probability.

Proof. Proof of (1) and (2).

Taking expectations in (5.19) and summing up (5.16), we have, for $1 \le m \le n$,

$$\mathbb{E} \log \Phi_{m,n}(1) = \sum_{k=n-m+1}^{n} \left[\Psi\left(\beta'(k-1) + 1 + \delta + \bar{\delta}\right) - \Psi\left(\beta'(k-1) + 1 + \bar{\delta}\right) \right] . \tag{2.10}$$

If in (2.10), we keep only the first two terms from the Abel-Plana formula (5.20), we get, for $m \le n-1$, or in the second regime, for $m = n$ large enough,

$$\begin{aligned}
&\int_{n-m}^{n} \left[\Psi\left(\beta'(s-1) + 1 + \delta + \bar{\delta}\right) - \Psi\left(\beta'(s-1) + 1 + \bar{\delta}\right) \right] ds \\
&+ \frac{1}{2} \Psi\left(\beta'(n-1) + 1 + \delta + \bar{\delta}\right) - \frac{1}{2} \Psi\left(\beta'(n-m-1) + 1 + \delta + \bar{\delta}\right) \\
&- \frac{1}{2} \Psi\left(\beta'(n-1) + 1 + \bar{\delta}\right) + \frac{1}{2} \Psi\left(\beta'(n-m-1) + 1 + \bar{\delta}\right) .
\end{aligned} \tag{2.11}$$

(Note that in the first regime and for $m = n$, we would get the terms $\Psi\left(-\beta' + 1 + \bar{\delta}\right)$ $\Psi\left(-\beta' + 1 + \delta + \bar{\delta}\right)$, which are not always well-defined). Integrating $\Psi = \ell'$, we get the following expression:

$$\begin{aligned}
L \quad := \quad & \widetilde{\ell}(\beta'(n-1) + \delta + \bar{\delta}) - \widetilde{\ell}(\beta'(n-m-1) + \delta + \bar{\delta}) \\
& - \widetilde{\ell}(\beta'(n-1) + \bar{\delta}) + \widetilde{\ell}(\beta'(n-m-1) + \bar{\delta}),
\end{aligned} \tag{2.12}$$

where $\widetilde{\ell}(x) := \frac{1}{\beta'} \ell(x+1) + \frac{1}{2} \Psi(x+1)$. From the Binet formula we have, for $x > -1$,

$$\widetilde{\ell}(x) = \frac{1}{\beta'} x \log x + \left(\frac{1}{\beta} + \frac{1}{2} \right) \log x - \frac{x-1}{\beta'} + r_1(x) \tag{2.13}$$

where

$$r_1(x) = \frac{1}{\beta'} \int_0^\infty f(s)[e^{-sx} - e^{-s}]\, ds + \frac{1}{2} \int_0^\infty e^{-sx} \left(\frac{1}{2} - s f(s)\, ds \right) . \tag{2.14}$$

Now, set

$$\begin{aligned}
I_1(n, m; \delta) \quad = \quad & \mathcal{I}\left(n - 1 + \frac{\delta + \bar{\delta}}{\beta'} \right) - \mathcal{I}\left(n - m - 1 + \frac{\delta + \bar{\delta}}{\beta'} \right) \\
& - \mathcal{I}\left(n - 1 + \frac{\bar{\delta}}{\beta'} \right) + \mathcal{I}\left(n - m - 1 + \frac{\bar{\delta}}{\beta'} \right)
\end{aligned} \tag{2.15}$$

and

$$I_2(n, m; \delta) = \left(\frac{1}{\beta} + \frac{1}{2} \right) J_2(n, m; \delta)$$

with

$$\begin{aligned}
J_2(n, m; \delta) \quad = \quad & \log\left(n - 1 + \frac{\delta + \bar{\delta}}{\beta'} \right) - \log\left(n - m - 1 + \frac{\delta + \bar{\delta}}{\beta'} \right) \\
& - \log\left(n - 1 + \frac{\bar{\delta}}{\beta'} \right) + \log\left(n - m - 1 + \frac{\bar{\delta}}{\beta'} \right) .
\end{aligned} \tag{2.16}$$

We have:

$$L = I_1(n, m; \delta) + I_2(n, m, \delta) + R\,,$$

where

$$\begin{aligned} R &= r_1(\beta'(n-1)+\delta+\bar{\delta}) - r_1(\beta'(n-m-1)+\delta+\bar{\delta}) \\ &\quad -r_1(\beta'(n-1)+\bar{\delta}) + r_1(\beta'(n-m-1)+\bar{\delta}). \end{aligned}$$

In the sequel, we use several times the trivial estimates (for c fixed and x tending to infinity)

$$\log(x+c) = \log x + o(1)\ ;\ \mathcal{I}(x+c) = \mathcal{I}(x) + c\log x + o(1)\,. \tag{2.17}$$

Let us set $m := \lfloor nt \rfloor$, and $t_n = m/n$, and let us suppose that we are in the first regime. If $0 < t < 1$, we have

$$\begin{aligned} I_1(n,m,\delta) &= \frac{\delta}{\beta'}\log\left(n-1+\frac{\bar{\delta}}{\beta'}\right) - \frac{\delta}{\beta'}\log\left(n-nt_n-1+\frac{\bar{\delta}}{\beta'}\right) + o(1) \\ &= -\frac{\delta}{\beta'}\log(1-t) + o(1), \end{aligned}$$

and

$$J_2(n,m,\delta) = o(1),$$

which implies

$$L = -\frac{\delta}{\beta'}\log(1-t) + R + o(1).$$

In the first regime and for $t = 1$, we have to estimate

$$\mathbb{E}\log\Phi_n(1) = \mathbb{E}\log\Phi_{n-1,n}(1) + \Psi\left(1+\delta+\bar{\delta}\right) - \Psi\left(1+\bar{\delta}\right).$$

Since the constant C can be modified, it is equivalent to deal with $m = n-1$ or with $m = n$. Taking $m = n-1$ gives for some constants C_1 and C_2,

$$\begin{aligned} I_1(n,n-1,\delta) &= \frac{\delta}{\beta'}\log\left(n-1+\frac{\bar{\delta}}{\beta'}\right) + \mathcal{I}\left(\frac{\delta+\bar{\delta}}{\beta'}\right) - \mathcal{I}\left(\frac{\bar{\delta}}{\beta'}\right) + o(1) \\ &= \frac{\delta}{\beta'}\log n + C_1 + o(1), \end{aligned}$$

$$J_2(n,n-1,\delta) = \log\left(\frac{\bar{\delta}}{\beta'}\right) - \log\left(\frac{\delta+\bar{\delta}}{\beta'}\right) + o(1) = C_2 + o(1),$$

which implies, for some constant C_3,

$$L = \frac{\delta}{\beta'}\log n + R + C_3 + o(1).$$

Let us now assume that we are in the second regime. For n large enough, we check the following estimates, uniform in $t \in (0,1]$:

$$I_1(n, nt_n, n\beta'\mathrm{d}) = n\mathcal{E}_\mathrm{d}(t_n) - J_2(n, nt_n, n\beta'\mathrm{d}) + o(1),$$

and

$$J_2(n, nt_n, n\beta'\mathrm{d}) = \mathcal{F}_\mathrm{d}(t_n) + o(1),$$

which implies

$$L = n\mathcal{E}_\mathrm{d}(t_n) + \left(\frac{1}{\beta} - \frac{1}{2}\right)\mathcal{F}_\mathrm{d}(t_n) + R + o(1).$$

Hence, (1) and (2) are proven, if we check that in any of the previous situations, $\mathbb{E}\log\Phi_{m,n}(1) - L$ and R tend to a constant when n goes to infinity, that this constant is zero, except perhaps in the first regime for $t = 1$, and that the convergence is uniform in t in the second regime. The first quantity can be expressed in function of the last integral term of the Abel-Plana formula, and the second one comes from the remaining integral term of the Binet formula.

First remaining term: Abel-Plana

We have to prove the convergence of:

$$\int_0^\infty \frac{g(n-m+\mathrm{i}y)-g(n+\mathrm{i}y)-g(n-m-\mathrm{i}y)+g(n-\mathrm{i}y)}{e^{2\pi y}-1}dy,$$

when n goes to infinity, for

$$g(z):=\Psi(\beta'(z-1)+1+\delta+\bar{\delta})-\Psi(\beta'(z-1)+1+\bar{\delta}),$$

the limit being zero, except in the first regime for $t=1$. Moreover, in the second regime, we want to check that the convergence is uniform with respect to t, i.e. with respect to $m\in\{1,\dots,n\}$. Note that in the second regime, the function g depends on n via $\delta=nd\beta'$. It is then sufficient to check the following result: for any sequence $(k_n)_{n\geq 1}$ of integers such that $1\leq k_n\leq n$,

$$\int_0^\infty \frac{g(k_n+\mathrm{i}y)-g(k_n-\mathrm{i}y)}{e^{2\pi y}-1}dy$$

is well-defined for all $n\geq 1$, and tends to zero when n goes to infinity, in any case and uniformly with respect to the sequence $(k_n)_{n\geq 1}$ in the second regime, and when $(k_n)_{n\geq 1}$ tends to infinity in the first regime. Indeed, we get the desired result by taking successively $k_n=n$ and $k_n=n-m$. Note that in the first regime for $t=1$, we need to take $k_n=1$ ($m=n-1$), which gives an integral independent of n, possibly different from zero. Now, we have

$$g(k_n\pm\mathrm{i}y)=\Psi(A'\pm\beta'\mathrm{i}y)-\Psi(A_0+\mathrm{i}B_0\pm\beta'\mathrm{i}y),$$

$A_0, A'>0$, $B_0\in\mathbb{R}$ being given by

$$A_0+\mathrm{i}B_0=\beta'(k_n-1)+1+\bar{\delta}$$

and

$$A'=\beta'(k_n-1)+1+\delta+\bar{\delta}.$$

In any case, A_0 and A' tend to infinity with n, since in the first regime, k_n goes to infinity, and in the second regime, it is the case for the real part of δ. Moreover, in the second regime, A and A' are greater than $n\beta'\,\mathfrak{Re}\,\mathrm{d}$, independently from the sequence $(k_n)_{n\geq 1}$. Hence, it is sufficient to check that

$$\int_0^\infty \frac{\Psi(A+\mathrm{i}B+\mathrm{i}C)-\Psi(A+\mathrm{i}B-\mathrm{i}C)}{e^{2\pi C/\beta'}-1}\,dC$$

is well-defined for $A>0$ and $B\in\mathbb{R}$, and tends to zero, uniformy in B, for $A\to\infty$. From (5.10) we have the following:

$$-\mathrm{i}\left(\Psi(A+\mathrm{i}B+\mathrm{i}C)-\Psi(A+\mathrm{i}B-\mathrm{i}C)\right)=2C\sum_{k=0}^\infty\frac{1}{(A+k+\mathrm{i}B)^2+C^2}\,.$$

To estimate the sum of this series, we first apply the crude estimates $|z|\geq|\mathfrak{Im}\,z|$ and $|z|\geq|\mathfrak{Re}\,z|$ to the denominator:

$$\begin{aligned}|(A+k+\mathrm{i}B)^2+C^2| &\geq 2|B||A+k|\\ |(A+k+\mathrm{i}B)^2+C^2| &\geq |(A+k)^2-B^2+C^2|.\end{aligned}$$

The first inequality implies, for $A+k \leq 2|B|$,

$$|(A+k+\mathrm{i}B)^2 + C^2| \geq (A+k)^2,$$

and the second implies, for $A+k > 2|B|$,

$$|(A+k+\mathrm{i}B)^2 + C^2| \geq \frac{3}{4}(A+k)^2 + C^2.$$

Hence,

$$\left|\sum_{k=0}^{\infty} \frac{1}{(A+k+\mathrm{i}B)^2 + C^2}\right| \leq \frac{4}{3}\sum_{k=0}^{\infty} \frac{1}{(A+k)^2} \leq \frac{4}{3}\left(\frac{1}{A^2} + \frac{1}{A}\right).$$

Therefore,

$$\int_0^\infty \frac{|\Psi(A+\mathrm{i}B+\mathrm{i}C) - \Psi(A+\mathrm{i}B-\mathrm{i}C)|}{e^{2\pi C/\beta'} - 1}\, dC \leq \frac{4}{3}\left(\frac{1}{A^2} + \frac{1}{A}\right)\int_0^\infty \frac{2C}{e^{2\pi C/\beta'} - 1} dC,$$

where the last integral is finite and does not depend on A and B.

Second remaining term: Binet

The terms involved in the expression of R are of the form:

$$R_1(x,y) = \int_0^\infty f(s)\left[e^{-sx} - e^{-sy}\right] ds \tag{2.18}$$

and

$$R_2(x,y) = \int_0^\infty \left(\frac{1}{2} - sf(s)\right)\left[e^{-sx} - e^{-sy}\right] ds, \tag{2.19}$$

where $x = \beta'(n-m-1) + \alpha$ and $y = \beta'(n-1) + \alpha$ with successively $\alpha = \delta + \bar{\delta}$ and $\alpha = \bar{\delta}$. If $t < 1$ or if we are in the second regime, the real parts of x and y tend to infinity with n: moreover, the convergence is uniform in t in the second regime. Then, as in [22], the dominated convergence theorem allows to conclude that $R_1(x,y)$ and $R_2(x,y)$ tend to zero, uniformly in t in the second regime. If $t = 1$ and if we are in the first regime, then y tends to infinity and $x = \alpha$ does not depend on n, which implies that $R_1(x,y)$ and $R_2(x,y)$ are still converging when n goes to infinity.

Proof of (3) and (4): Computation of covariances.

By independence of the variables γ_j, we can sum up variances or covariances issued from (5.18). We can then prove the announced result in the same way as (1) and (2), using the Abel-Plana summation and the Binet formula again, with Ψ replaced by Ψ': we omit the detail.

Proof of (5): the process $(\zeta_n(t))_{0\leq t\leq 1}$ is a two-dimensional martingale. Hence, by Doob's inequality:

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,1]} |\zeta_n(t)|^2\right] \leq 4\mathbb{E}[|\zeta_n(1)|^2] &= 4\,\mathrm{Tr}(\mathrm{Cov}\,\zeta_n(1)) \\ \underset{n\to\infty}{\longrightarrow} \frac{1}{\beta'}\int_0^1 \frac{dt}{1-t+2\mathfrak{Re}\,\mathrm{d}} &= \frac{1}{\beta'}\log\left(\frac{1+2\mathfrak{Re}\,\mathrm{d}}{2\mathfrak{Re}\,\mathrm{d}}\right) < \infty.\end{aligned}$$

Moreover, by the uniform convergence (2),

$$\sup_{t\in[0,1]}\left|\mathbb{E}\log\Phi_{\lfloor nt\rfloor,n}(1) - n\mathcal{E}_{\mathrm{d}}\left(\frac{\lfloor nt\rfloor}{n}\right)\right| \underset{n\to\infty}{\longrightarrow} \left(\frac{1}{2} - \frac{1}{\beta}\right)\sup_{t\in[0,1]}|\mathcal{F}_{\mathrm{d}}(t)| < \infty,$$

and one also has

$$\sup_{t\in[0,1]} n\left|\mathcal{E}_{\mathrm{d}}\left(\frac{\lfloor nt\rfloor}{n}\right) - \mathcal{E}_{\mathrm{d}}(t)\right| \leq \sup_{t\in[0,1]} |\mathcal{E}_{\mathrm{d}}'(t)| < \infty.$$

Combining these estimates gives the following L^2 bound:

$$\sup_{n\geq 1} \mathbb{E}\left[\left(\sup_{t\in[0,1]} \left|\log \Phi_{\lfloor nt\rfloor,n}(1) - n\mathcal{E}_{\mathrm{d}}(t)\right|\right)^2\right] < \infty,$$

which gives (4) after dividing by n^2.

2.2. **Fluctuations.** We now state a theorem about some limiting distributions related to the process $\{\Phi_{\lfloor nt\rfloor,n}(1), t \in [0,1]\}$. Since it can be shown by adapting the arguments of Killip and Stoiciu [17] for similar results, we omit the proof of our theorem, except for the last part.

Let D_T and D be the space of càdlàg $\mathbb{R}^2$-valued functions on $[0,T]$ and $[0,1)$ respectively, starting from zero. The next theorem is about some limiting distributions related to the process $\{\zeta_n(t), t \in [0,1]\}$ or $\{\zeta_n(t), t < 1\}$. More precisely let $(\mathcal{Z}_t^{\mathrm{d}})^{1/2}$ denote the positive symmetric square root of $\mathcal{Z}_t^{\mathrm{d}}$ defined in (2.8) and let $\mathbf{B}_t$ be a standard two dimensional Brownian motion.

Theorem 2.2. (1) *If $\delta = \beta'\mathrm{d}n$ with $\mathfrak{Re}\,\mathrm{d} > 0$ (second regime), then as $n \to \infty$ the process $\{\zeta_n(t)\ ;\ t \in [0,1]\}_n$ converges in distribution in the Skorokhod space D_1 to the Gaussian diffusion $\{\zeta_t^{\mathrm{d}}\ ;\ t \in [0,1]\}$, solution of the stochastic differential equation:*

$$d\zeta_t^{\mathrm{d}} = (\mathcal{Z}_t^{\mathrm{d}})^{1/2} d\mathbf{B}_t\,. \tag{2.20}$$

(2) *If δ is fixed with $\mathfrak{Re}\,\delta > -1/2$ (first regime), then the joint law of the process $\{\zeta_n(t)\ ;\ t < 1\}$ (with trajectories in the Skorokhod space D) and the variable*

$$\Theta := \frac{\log \Phi_n(1) - \frac{\delta}{\beta'}\log n}{\sqrt{\log n}}$$

converges, when n goes to infinity, to the joint distribution of $\{\zeta_t^0\ ;\ t < 1\}$ and $\mathcal{N}(0;\beta^{-1}) + \mathrm{i}\,\mathcal{N}(0;\beta^{-1})$, the process and the two gaussian variables being independent.

Notice that the convergence in law of Θ is an extension of the celebrated Keating and Snaith result [14].

Proof. In order to prove (1), and in (2), the convergence of $\{\zeta_n(t)\ ;\ t < 1\}$ and Θ, taken separately, we apply a version of the Lindeberg-Lévy-Lyapunov criterion, available for convergence of processes ([13] Chap. 3c). For $t < 1$, or in the second regime, it is enough to prove that

$$\sum_{k=0}^{\lfloor nt\rfloor-1} \mathbb{E}|A_k|^4 \underset{n\to\infty}{\longrightarrow} 0, \tag{2.21}$$

where

$$A_k = \log(1-\gamma_k) - \mathbb{E}[\log(1-\gamma_k)].$$

For $t = 1$ in the first regime, it is sufficient to check:

$$\frac{1}{\log^2 n} \sum_{k=0}^{n-1} \mathbb{E}|A_k|^4 \underset{n\to\infty}{\longrightarrow} 0. \tag{2.22}$$

Now,

$$\begin{aligned}
\mathbb{E}|A_k|^4 &\leq 8\mathbb{E}(\mathfrak{Re}\, A_k)^4 + 8\mathbb{E}(\mathfrak{Im}\, A_k)^4 \\
&\leq 24 \operatorname{Var}^2(\mathfrak{Re}\, \log(1-\gamma_k)) + 24 \operatorname{Var}^2(\mathfrak{Re}\, \log(1-\gamma_k)) \\
&\quad + 8\kappa_4(\mathfrak{Re} \log(1-\gamma_k)) + 8\kappa_4(\mathfrak{Im} \log(1-\gamma_k)),
\end{aligned}$$

where κ_4 denotes the fourth cumulant. Using the second and fourth order derivatives of the function Λ introduced in the appendix, we see that $\mathbb{E}|A_k|^4$ is a linear combination of terms of the form $\Psi'''(r+1+\alpha)$ and $\Psi'(r+1+\alpha)\Psi'(r+1+\alpha')$, for $r = \beta'(n-k-1)$ and $\alpha, \alpha' \in \{\delta, \bar{\delta}, \delta + \bar{\delta}\}$. Now, by (5.9), for $\mathfrak{Re}\, x > 0$,

$$\Psi'(x) = \frac{1}{x} + O\left(\frac{1}{(\mathfrak{Re}\, x)^2}\right)$$

and

$$\Psi'''(x) = \frac{2}{x^3} + O\left(\frac{1}{(\mathfrak{Re}\, x)^4}\right).$$

Hence, all the terms involved in the expression of $\mathbb{E}|A_k|^4$ are dominated by $(n-k)^{-2}$, and they are dominated by n^{-2} in the second regime or for $t < 1$. Hence,

$$\sum_{k=0}^{\lfloor nt \rfloor - 1} \mathbb{E}|A_k|^4$$

is dominated by $1/n$, except in the first regime for $t = 1$, in which case it is bounded. This shows the desired results (2.21) and (2.22).

In order to prove the convergence of the joint distribution in (2), we can follow the scheme of [22] p.209. We take $0 < t_0 < t_1 < 1$. From the above results,

$$\frac{\log \Phi_{\lfloor nt_1 \rfloor, n}(1)}{\sqrt{\log n}} \to 0 \tag{2.23}$$

in probability, so that

$$\frac{\log \Phi_n(1) - \log \Phi_{\lfloor nt_1 \rfloor, n}(1) - \frac{\delta}{\beta'} \log n}{\sqrt{\log n}} \to \mathcal{N}(0; \beta^{-1}) + \mathrm{i}\mathcal{N}(0; \beta^{-1}),$$

the two gaussian variables being independent. Now, for n large enough,

$$\log \Phi_n(1) - \log \Phi_{\lfloor nt_1 \rfloor, n}(1) = \sum_{k=\lfloor nt_1 \rfloor}^{n-1} (1-\gamma_k),$$

is independent of $\{\zeta_n(t)\ ;\ t \leq t_0\}$, which is function of $(\gamma_k)_{k<nt_0}$ (recall that the variables γ_k are independent). Since, $\{\zeta_n(t)\ ;\ t \leq t_0\}$ tends in law to $\{\zeta_t^0\ ;\ t \leq t_0\}$ (as a process with trajectories in D_{t_0}), we deduce that

$$\left(\{\zeta_n(t)\ ;\ t \leq t_0\}; \frac{\log \Phi_n(1) - \log \Phi_{\lfloor nt_1 \rfloor, n}(1) - \frac{\delta}{\beta'} \log n}{\sqrt{\log n}}\right),$$

tends in law to

$$\left(\{\zeta_t^0\ ;\ t \leq t_0\}; \mathcal{N}(0; \beta^{-1}) + \mathrm{i}\mathcal{N}(0; \beta^{-1})\right),$$

$\{\zeta_t^0\ ;\ t\le t_0\}$ being independent of $\mathcal{N}(0;\beta^{-1})+\mathrm{i}\mathcal{N}(0;\beta^{-1})$. Using again (2.23), we deduce that

$$\left(\{\zeta_n(t)\ ;\ t\le t_0\};\frac{\log\Phi_n(1)-\frac{\delta}{\beta'}\log n}{\sqrt{\log n}}\right)$$

has the same limiting distribution. Taking $t_0\to 1$ gives part (2) of the theorem.

3. Large deviations

3.1. **Notation and main statements.** Throughout this section, we use the standard notation of [7]. In particular we write LDP for Large Deviation Principle. We say that a sequence $(P_n)_{n\in\mathbb{N}}$ of probability measures on a Polish space $\mathcal{X}$ satisfies a LDP with speed a_n (going to infinity with n) and rate function I iff $I:\mathcal{X}\to[0,\infty]$ is lower semicontinuous and if

$$\text{For every open set }\ O\subset\mathcal{X},\ \liminf_n\frac{1}{a_n}\log P_n(O)\ge-\inf_O I\,,$$

$$\text{For every closed set }\ F\subset\mathcal{X},\ \liminf_n\frac{1}{a_n}\log P_n(F)\ge-\inf_F I\,.$$

The rate function is good if its level sets are compact. Moreover, if X_n are $\mathcal{X}$ random variables distributed according to P_n, we say that the sequence (X_n) satisfies the LDP if the sequence (P_n) satisfies the LDP.

The reader may have some interest in consulting [8] and mainly [22] where a similar method is used for a different model.

For $T\le 1$, let M_T stand for the space of signed measures on $[0,T]$ and let $\mathrm{M}_<$ be the subspace of M_1 consisting of measures whose support is a compact subset of $[0,1)$. We endow $D\times D$ with the weak topology $\sigma(D\times D,\mathrm{M}_<\times\mathrm{M}_<)$. So, $D\times D$ is the projective limit of the family, indexed by $T<1$, of the topological spaces $(D_T\times D_T,\sigma(D_T\times D_T,\mathrm{M}_T\times\mathrm{M}_T))$.

Let V_ℓ (resp. V_r) be the space of left (resp. right) continuous $\mathbb{R}$-valued functions with bounded variations. We put a superscript T to specify the functions on $[0,T]$. There is a bijective correspondence between V_r^T and M_T:

- for any $v\in V_r^T$, there exists a unique $\mu\in\mathrm{M}_T$ such that $v=\mu([0,\cdot])$; we denote it by $\dot v$,

- for any $\mu\in\mathrm{M}_T$, $v=\mu([0,\cdot])$ is in V_r^T.

For $v\in D$, let $\dot v=\dot v_a+\dot v_s$ be the Lebesgue decomposition of the measure $\dot v$ in absolutely continuous and singular parts with respect to the (vectorial) Lebesgue measure. The measure $\dot v_a$ can then be identified with its density.

Now, for $\xi,\eta\in\mathbb{R}$, let us define:

$$H_a(\xi,\eta):=-\xi-\log(2\cos\eta-e^\xi)\,, \tag{3.1}$$

if $|\eta|<\pi/2$ and $2\cos\eta-e^\xi>0$, and $H_a(\xi,\eta)=\infty$ otherwise. For $\varphi,\psi\in V_r$ and $T\le 1$, let us denote:

$$\mathcal{H}_0(T,\varphi,\psi):=\begin{cases}\int_0^T(1-\tau)H_a(\dot\varphi_a(\tau),\dot\psi_a(\tau))\,d\tau+\int_0^T(1-\tau)d(-\dot\varphi_s)(\tau)\\ \text{if } \dot\psi_s=0 \text{ and } -\dot\varphi_s \ \text{ is a positive measure}\,,\\ +\infty\quad\text{otherwise.}\end{cases} \tag{3.2}$$

Theorem 3.1. (1) *For every $T<1$, the sequence*

$$\{n^{-1}\left(\mathfrak{Re}\log\Phi_{\lfloor nt\rfloor,n}(1),\mathfrak{Im}\log\Phi_{\lfloor nt\rfloor,n}(1)\right)\ ;\ t\in[0,T]\}$$

under the $C\beta E^{(n)}$ *measure (i.e. first regime and* $\delta = 0$*) satisfies in* $(D_T \times D_T, \sigma(D_T \times D_T, \mathrm{M}_T \times \mathrm{M}_T))$ *the LDP with speed* $\beta' n^2$ *with good rate function* $\mathcal{H}_0(T, \varphi, \psi)$*.*

(2) *For* $\mathfrak{Re}\,\mathrm{d} > 0$ *and* $T \le 1$*, the sequence*

$$\{n^{-1}\left(\mathfrak{Re} \log \Phi_{\lfloor nt \rfloor, n}(1), \mathfrak{Im} \log \Phi_{\lfloor nt \rfloor, n}(1)\right) \ t \in [0, T]\}$$

under the $\mathrm{CJ}^{(n)}_{\beta, \beta' \mathrm{d} n}$ *measure (second regime) satisfies in* $(D_T \times D_T, \sigma(D_T \times D_T, \mathrm{M}_T \times \mathrm{M}_T))$ *the LDP with speed* $\beta' n^2$ *with good rate function*

$$\mathcal{H}_{\mathrm{d}}(T, \varphi, \psi) = \mathcal{H}_0(T, \varphi, \psi) - 2(\mathfrak{Re}\,\mathrm{d})\varphi(T) - 2(\mathfrak{Im}\,\mathrm{d})\psi(T) - \mathcal{C}_{\mathrm{d}}(T) \tag{3.3}$$

with

$$\begin{aligned} \mathcal{C}_{\mathrm{d}}(T) &= F(1 + \mathrm{d}) - F(1 - T + \mathrm{d}) + F(1 + \bar{\mathrm{d}}) - F(1 - T + \bar{\mathrm{d}}) \\ &\quad - F(1 + 2\mathfrak{Re}\,\mathrm{d}) + F(1 - T + 2\mathfrak{Re}\,\mathrm{d}) - F(1) + F(1 - T)\,. \end{aligned}$$

We give now a result on the marginal at time T fixed. It is obtained by applying the contraction principle to the mapping

$$(\varphi, \psi) \mapsto (\varphi(T), \psi(T))\,.$$

In all the sequel of the paper, we consider either the $C\beta E^{(n)}$ ensemble (first regime for $\delta = 0$), or the $\mathrm{CJ}^{(n)}_{\beta, \beta' \mathrm{d} n}$ ensemble for $\mathfrak{Re}\,\mathrm{d} > 0$ (second regime). In the first case, we put $\mathrm{d} = 0$.

Theorem 3.2. *When* $(\mathfrak{Re}\,\mathrm{d} > 0, T \le 1)$ *or* $(\mathrm{d} = 0, T < 1)$*, the sequence*

$$\{n^{-1}\left(\mathfrak{Re} \log \Phi_{\lfloor nT \rfloor}(1), \mathfrak{Im} \log \Phi_{\lfloor nT \rfloor}(1)\right)\}_n$$

satisfies the LDP in $\mathbb{R}^2$ *with speed* $\beta' n^2$ *with good rate function*

$$h_{\mathrm{d}}(T, \xi, \eta) = \inf\{\mathcal{H}_{\mathrm{d}}(T, \varphi, \psi) \mid \varphi(T) = \xi, \psi(T) = \eta\}\,. \tag{3.4}$$

In particular (cf. (3.3))

$$h_{\mathrm{d}}(T, \xi, \eta) = h_0(T, \xi, \eta) - 2(\mathfrak{Re}\,\mathrm{d})\xi - 2(\mathfrak{Im}\,\mathrm{d})\eta - \mathcal{C}_{\mathrm{d}}(T), \tag{3.5}$$

which allows to compute easily h_{d} *when* h_0 *is known.*

For the two coordinates, separately, we have the following known result, which comes from formula (C.5) in [12].

Theorem 3.3. *Assume* $\mathrm{d} = 0$ *and* $\beta = 2$*.*

(1) *The sequence* $\{n^{-1}\,\mathfrak{Re} \log \Phi_n(1)\}_n$ *satisfies the LDP in* $\mathbb{R}$ *with speed* n^2 *and rate function given by the dual (Legendre) of the function:*

$$s \mapsto \frac{(1+s)^2}{2} \log(1+s) - \left(1 + \frac{s}{2}\right)^2 \log\left(1 + \frac{s}{2}\right) - \frac{s^2}{4} \log(2s)$$

for $s \ge 0$*, and by* ∞ *for* $s < 0$*. It vanishes for negative values of the argument, and it is infinite beyond* $\log 2$*.*

(2) *The sequence* $\{n^{-1}\,\mathfrak{Im} \log \Phi_n(1)\}_n$ *satisfies the LDP in* $\mathbb{R}$ *with speed* n^2 *and good rate function given by the dual (Legendre) of the function given by*

$$t \mapsto \frac{t^2}{8} \log\left(1 + \frac{4}{t^2}\right) - \frac{1}{2} \log\left(1 + \frac{t^2}{4}\right) + t \arctan(t/2)\,.$$

It is finite on $(-\pi/2, \pi/2)$*, and infinite otherwise.*

We now give the precise behavior of the first coordinate, in the particular case where d is real.

Theorem 3.4. (1) *For* $(\mathrm{d} = 0, T < 1)$*, or* $(\mathrm{d} > 0, T \le 1)$*, the sequence* $\{n^{-1}\,\mathfrak{Re}\,\log\Phi_n(T)\}_n$ *satisfies the LDP in* $\mathbb{R}$ *with speed* $\beta' n^2$ *with good rate function* $h_{\mathrm{d}}(T,\cdot,0)$*.*

(2) *Let* $\xi_T := \mathcal{J}(T) - 1 - \mathcal{J}\left(\frac{1+T}{2}\right) + \mathcal{J}\left(\frac{1-T}{2}\right) \le 0$*.*

(a) *If* $\xi \in [\xi_T, T\log 2)$ *the equation*

$$\mathcal{J}(1+\gamma) - \mathcal{J}(1-T+\gamma) - \mathcal{J}\left(1+\frac{\gamma}{2}\right) + \mathcal{J}\left(1-T+\frac{\gamma}{2}\right) = \xi \tag{3.6}$$

has a unique solution γ *and we have*

$$h_0(T,\xi,0) = \gamma\xi - \mathcal{L}_0(T,\gamma,0) \tag{3.7}$$

where

$$\begin{aligned} \mathcal{L}_0(T,\gamma,0) \quad &:= \quad F(1+\gamma) - F(1-T+\gamma) - F(1-T) + F(1) \\ &\qquad -2F\left(1+\frac{\gamma}{2}\right) + 2F\left(1-T+\frac{\gamma}{2}\right) \end{aligned} \tag{3.8}$$

(b) *If* $\xi < \xi_T$*, then*

$$h_0(T,\xi,0) = h_0(T,\xi_T,0) + (1-T)(\xi_T - \xi)\,. \tag{3.9}$$

(c) *If* $\xi \ge T\log 2$*, then* $h_0(T,\xi,0) = \infty$*.*

(3) *For* $(\mathrm{d} > 0, T \le 1)$ *or* $(\mathrm{d} = 0, T < 1)$*, the function* $\xi \mapsto h_{\mathrm{d}}(T,\xi,0)$ *is the dual function of*

$$\gamma \mapsto \mathcal{L}_{\mathrm{d}}(T,\gamma,0) := \mathcal{L}_0(T,\gamma+2\mathrm{d},0) - \mathcal{L}_0(T,2\mathrm{d},0)\,.$$

3.2. **Proof of Theorem 3.1.** We compute the normalized cumulant generating function, find its limit, perform the dual transform, study the exponential tightness and eventually prove that the LDP is satisfied.

From (5.12) and (5.13) we see that

$$\frac{\mathrm{CJ}_{\beta,\delta}\left[(1-\bar\gamma_j)^z(1-\gamma_j)^{\bar z}\right]}{\mathrm{CJ}_{\beta,0}\left[(1-\bar\gamma_j)^{(z+\delta)}(1-\gamma_j)^{(\bar z+\bar\delta)}\right]} = \frac{c_{r,\delta}}{c_{r,0}} \tag{3.10}$$

for every $j < n-1$ and z such that $2\mathfrak{Re}\,(\delta+z)) > -1$, where $r = \beta'(n-j-1)$. The RHS of (3.10) does not depend on z. It should then be clear that we can reduce the case $\mathfrak{Re}\,\mathrm{d} > 0$ to the case $\mathrm{d} = 0$. The above shift in the argument of the generating function provides the linear term $-2(\mathfrak{Re}\,\mathrm{d})\varphi(T) - 2(\mathfrak{Im}\,\mathrm{d})\psi(T)$ in the rate function, and the RHS of (3.10) gives the constant $-\mathcal{C}_{\mathrm{d}}(T)$.

• *First step: The normalized cumulant generating function.*

Lemma 3.5. *Let* $T < 1$*.*

(1) *Let us assume* $\mathrm{d} = 0$*.*

(a) *For every path* $(x(\tau), y(\tau))_{\tau\in[0,T]} \in V_\ell^T \times V_\ell^T$ *such that* $x(\tau)+1-\tau > 0$ *on* $[0,T]$*, set* $2z(\tau) := x(\tau) + \mathrm{i}y(\tau)$ *and*

$$\Lambda_0(T,x,y) := \int_0^T \left(\mathcal{J}(1-\tau+x(\tau)) - \mathcal{J}(1-\tau+z(\tau)) - \mathcal{J}(1-\tau+\bar z(\tau)) + \mathcal{J}(1-\tau)\right)d\tau\,. \tag{3.11}$$

Then we have:

(3.12)

$$\lim_{n\to\infty} \frac{1}{\beta' n^2} \log \mathbb{E} \exp\left(n\beta' \mathfrak{Re} \left[\int_0^T (x(\tau) - \mathrm{i} y(\tau))\, d\log \Phi_{\lfloor n\tau \rfloor}(1) \right] \right) = \Lambda_0(T, x, y)$$

(b) *In particular, when* $x(.) \equiv s$ *and* $y(.) \equiv t$*, if we set* $\mathcal{L}_0(T, s, t) := \Lambda_0(T, x, y)$ *(which generalizes the notation* $\mathcal{L}_0(T, \gamma, 0)$ *introduced in Theorem 3.4), we have for every* $s > -(1-T)$

$$\begin{aligned} \mathcal{L}_0(T,s,t) &= F(1+s) - F(1-T+s) + F(1) - F(1-T) \\ &\quad -F(1+z) + F(1-T+z) - F(1+\bar{z}) + F(1-T+\bar{z}) \end{aligned} \tag{3.13}$$

where $2z = s + \mathrm{i}t$.

(2) *For* $\mathfrak{Re}\,\mathrm{d} > 0$*, the analogues of (3.11) and (3.13) are*

$$\Lambda_{\mathrm{d}}(T, x, y) = \Lambda_0(T, x + 2\mathfrak{Re}\,\mathrm{d}, y + 2\mathfrak{Im}\,\mathrm{d}) - \Lambda_0(T, 2\mathfrak{Re}\,\mathrm{d}, 2\mathfrak{Im}\,\mathrm{d})\,, \tag{3.14}$$

and

$$\mathcal{L}_{\mathrm{d}}(T, s, t) = \mathcal{L}_0(T, s + 2\mathfrak{Re}\,\mathrm{d}, t + 2\mathfrak{Im}\,\mathrm{d}) - \mathcal{L}_0(T, 2\mathfrak{Re}\,\mathrm{d}, 2\mathfrak{Im}\,\mathrm{d})\,, \tag{3.15}$$

for $s > -(1-T) - 2\mathfrak{Re}\,\mathrm{d}$.

Proof. (1)(a) Let us set $x(\tau) = y(\tau) = z(\tau) := 0$ for $\tau > T$, and $\tau_j := (j+1)/n$ for $j = 0, \dots, n-1$. One has:

$$\begin{aligned} &\mathbb{E} \exp\left(n\beta' \mathfrak{Re} \left[\int_0^T (x(\tau) - \mathrm{i} y(\tau))\, d\log \Phi_{\lfloor n\tau \rfloor}(1) \right] \right) \\ &= \mathbb{E} \exp\left(n\beta' \mathfrak{Re} \left[\sum_{j=0}^{n-1} (x(\tau_j) - \mathrm{i} y(\tau_j))\, \log(1 - \gamma_j) \right] \right) \\ &= \prod_{j=0}^{n-1} \mathbb{E} \exp\left(n\beta' \mathfrak{Re} \left[(x(\tau_j) - \mathrm{i} y(\tau_j))\, \log(1 - \gamma_j) \right] \right) \end{aligned}$$

and using (5.14), we get:

$$\begin{aligned} &\log \mathbb{E} \exp\left(n\beta' \mathfrak{Re} \left[\int_0^T (x(\tau) - \mathrm{i} y(\tau))\, d\log \Phi_{\lfloor n\tau \rfloor}(1) \right] \right) \\ &= \sum_{j=0}^{n-1} \left[\ell(n\beta'(1 - \tau_j + x(\tau_j)) + 1) + \ell(n\beta'(1 - \tau_j) + 1) \right. \\ &\quad \left. -\ell(n\beta'(1 - \tau_j + z(\tau_j)) + 1) - \ell(n\beta'(1 - \tau_j + \bar{z}(\tau_j)) + 1) \right]. \end{aligned} \tag{3.16}$$

Now, the Binet formula (5.1) yields:

$$\ell(u+1) = \log(u) + \log\Gamma(u) = (u + 1/2)\log u - u + 1 + \int_0^\infty f(s)[e^{-su} - e^{-s}]ds\,.$$

If we apply four times this formula in order to estimate the term indexed by $j < n$ in (3.16), the contribution of the term u is 0, the contribution of the term $\log u$ is

$$\log(1 - \tau_j) + \log(1 - \tau_j + x(\tau_j)) - \log(1 - \tau_j + z(\tau_j)) - \log(1 - \tau_j + \overline{z(\tau_j)})\,,$$

the contribution of the term $u \log u$ is proportional to $\beta' n$ with coefficient

$$\mathcal{J}(1-\tau_j) - \mathcal{J}(1-\tau_j + z(\tau_j)) - \mathcal{J}(1-\tau_j + \overline{z(\tau_j)}) + \mathcal{J}(1-\tau_j + x(\tau_j))\,;$$

dividing by $\beta' n^2$ and performing Riemann sums gives the integral in (3.11). The remaining part is a sum of bounded terms which is negligible with respect to n^2.

(1)(b) The equality (3.13) is obvious by integration.

(2) To get the expression corresponding to $\mathfrak{Re}\,\mathrm{d} > 0$ we just use (3.10). ■

• *Second step ; the Legendre duality.*

It will be convenient to perform a time-change in (3.11), setting

$$x(\tau) = (1-\tau)X(\tau)\ ;\ y(\tau) = (1-\tau)Y(\tau)\ ;\ z(\tau) = (1-\tau)Z(\tau)\,. \tag{3.17}$$

The above expression (3.11)) of Λ_0 becomes

$$\Lambda_0(T, x, y) = \int_0^T (1-\tau) L(X(\tau), Y(\tau)) d\tau \tag{3.18}$$

where

$$L(X, Y) = \mathcal{J}(1+X) - \mathcal{J}(1+Z) - \mathcal{J}(1+\bar{Z}) \tag{3.19}$$

i.e.

$$\begin{aligned} L(X,Y) &= (1+X)\log(1+X) + Y \arctan\frac{Y}{2+X} \\ &\quad - \left(1+\frac{X}{2}\right)\log\left[\left(1+\frac{X}{2}\right)^2 + \frac{Y^2}{4}\right]. \end{aligned} \tag{3.20}$$

Looking for the Legendre dual, we see that for $|\eta| < \pi/2$ and $e^{\xi} < 2\cos\eta$, the supremum

$$L^\star(\xi,\eta) = \sup_{X,Y} X\xi + Y\eta - L(X,Y)$$

is achieved in (X, Y) satisfying

$$\frac{1+X}{\sqrt{\left(1+\frac{X}{2}\right)^2 + \frac{Y^2}{4}}} = e^{\xi}\ ,\ \frac{Y}{2+X} = \tan\eta \tag{3.21}$$

i.e.

$$X = \frac{e^{\xi} - \cos\eta}{\cos\eta - \frac{1}{2}e^{\xi}}\ ,\ Y = \frac{\sin\eta}{\cos\eta - \frac{1}{2}e^{\xi}}\,.$$

Note that under the assumption above, X is admissible, i.e. $X > -1$. One deduces that (cf. (3.1))

$$L^\star(\xi,\eta) = -\xi - \log(2\cos\eta - e^{\xi}) = H_a(\xi,\eta). \tag{3.22}$$

On the other hand, one can check that $L^\star(\xi,\eta)$ is infinite if $|\eta| \geq \pi/2$ or $e^{\xi} \geq 2\cos\eta$.

One deduces that there exists a recession function:

$$(\xi,\eta) \mapsto \lim_{\kappa\to+\infty} \kappa^{-1} L^\star(\kappa\xi,\kappa\eta) = \begin{cases} -\xi \text{ if } \xi < 0 \text{ and } \eta = 0\,, \\ \infty \text{ otherwise}\,. \end{cases}$$

This function can be used to obtain the rate function $\mathcal{H}_0(T,\varphi,\psi)$ when the measures $\dot\varphi_s$ and $\dot\psi_s$ are not zero. By using the same methods as in [22], one deduces

$$\begin{aligned}\mathcal{H}_0(T,\varphi,\psi) &= \sup_{x(.),y(.)}\left[\int_0^T \Big(x(\tau)d\dot\varphi(\tau)+y(\tau)d\dot\psi(\tau)\Big) - \Lambda_0(T,x(.),y(.))\right]\\ &= \sup_{X(.),Y(.)}\int_0^T (1-\tau)\Big[X(\tau)d\dot\varphi(\tau)+Y(\tau)d\dot\psi(\tau) - L(X(\tau),Y(\tau))\,d\tau\Big]\\ &= \int_0^T (1-\tau)H_a(\dot\varphi_a(\tau),\dot\psi_a(\tau))\,d\tau + \int_0^T (1-\tau)d(-\dot\varphi_s)(\tau)\,,\end{aligned}$$

where the second equality comes from (3.17) and (3.18), and where the last equality comes from [21] Theorem 5.

The value of the constant $\mathcal{C}_{\mathrm{d}}(T)$ is obtained owing to (3.13):

$$\mathcal{C}_{\mathrm{d}}(T) = -\mathcal{L}_0(T, 2\mathfrak{Re}\,\mathrm{d}, 2\mathfrak{Im}\,\mathrm{d})\,.$$

• *Third step : exponential tightness.*

Exponential tightness is not needed for the second argument since it lives in $[-\pi/2,\pi/2]$. For the first, we have $|\log x| \le -\log x + 2\log 2$ for $x \le 2$, hence:

$$\mathbb{P}_{\mathrm{d}}(\sum_{j\le nT-1} |\log|1-y_j|| \ge na) \le \mathbb{P}_{\mathrm{d}}(\sum_{j\le nT-1} -\log(1-y_j) \ge n(a-2T\log 2)).$$

Now, for $\theta < 0$ (by Chernov inequality),

$$\mathbb{P}_{\mathrm{d}}(\sum_{j\le nT-1} -\log(1-y_j) \ge n(a-2T\log 2)) \le e^{n^2\beta'\theta(a-2T\log 2)}\mathbb{E}_{\mathrm{d}}\left(|\Phi_{\lfloor nT\rfloor}(1)|^{n^2\beta'\theta}\right)$$

so that, taking logarithm and applying (3.13) we get, for $\theta \in (-(1-T)-2\mathfrak{Re}\,\mathrm{d}, 0)$

$$\limsup_{n\to\infty}(\beta' n^2)^{-1}\log \mathbb{P}_{\mathrm{d}}(\sum_{j\le nT} |\log|1-y_j| \ge na) \le \theta(a-2T\log 2) + \mathcal{L}_{\mathrm{d}}(T,\theta,0).$$

It remains to let $a\to\infty$ to get the exponential tightness. We remark that when $\mathrm{d}=0$, the exponential tightness holds only for $T<1$. ■

Remark 3.6. *The mean trajectory is obtained when $H_a(\dot\varphi,\dot\psi)\equiv 0$ i.e. $\cos\dot\psi = \cosh\dot\varphi$ or $\dot\varphi=\dot\psi=0$*

3.3. **Comment on Theorem 3.2.** Let us study the variational problem (3.4) issued from the contraction. Using (3.2) and (3.22), we see that the Euler equation is

$$\begin{aligned}&\frac{d}{d\tau}\left((1-\tau)\frac{\partial L^\star}{\partial\dot\varphi}\right) = 0\\ &\frac{d}{d\tau}\left((1-\tau)\frac{\partial L^\star}{\partial\dot\psi}\right) = 0\,,\end{aligned}\tag{3.23}$$

and the optimal path is then given by

$$\dot\varphi(\tau) = \frac{\partial L}{\partial X}\left(\frac{\gamma}{1-\tau},\frac{\rho}{1-\tau}\right)\ ,\ \dot\psi(\tau) = \frac{\partial L}{\partial Y}\left(\frac{\gamma}{1-\tau},\frac{\rho}{1-\tau}\right)$$

i.e.

$$\dot\varphi(\tau) = \log(1-\tau+\gamma) - \frac{1}{2}\left(\log\left(1-\tau+\frac{\gamma}{2}\right)^2 + \frac{\rho^2}{4}\right), \tag{3.24}$$

$$\dot\psi(\tau) = \arctan\frac{\rho}{2(1-\tau)+\gamma}. \tag{3.25}$$

This path will be admissible if there exist γ and ρ such that

$$\int_0^T \dot\varphi(\tau)d\tau = \xi\ ,\ \int_0^T \dot\psi(\tau)d\tau = \eta\,. \tag{3.26}$$

When the path is admissible, we have

$$H_a(\dot\varphi(\tau),\dot\psi(\tau)) = \widehat{\mathcal{L}}(\dot\varphi(\tau),\dot\psi(\tau)) = \frac{\gamma}{1-\tau}\dot\varphi(\tau) + \frac{\rho}{1-\tau}\dot\psi(\tau) - L\left(\frac{\gamma}{1-\tau},\frac{\rho}{1-\tau}\right)$$

and

$$h_0(T,\xi,\eta) = \gamma\xi + \rho\eta - \int_0^T (1-\tau)L\left(\frac{\gamma}{1-\tau},\frac{\rho}{1-\tau}\right)d\tau.$$

3.4. **Comment on Theorem 3.4.** In the case $\mathrm{d}=0$, $T=1$, [12] proved the LDP by tackling directly the normalized cumulant generating function. This gives an incomplete LDP since there is no steepness in 0. They use a Fourier inversion to take into account the negative side. We see that the function $\mathcal{L}_0(1,s,0)$ is the limiting n.c.g.f. of $\log|\Phi_n(1)|$, computed in [12] Theorem 3.3. Besides, $\mathcal{L}_0(1,0,t)$ is the limiting n.c.g.f. of the argument of $\Phi_n(1)$, computed in [12], Theorem 3.4, and

$$\lim_{t\to\pm\infty}\frac{\mathcal{L}_0(1,0,t)}{t} = \pm\frac{\pi}{2}\,.$$

3.5. **Proof of Theorem 3.4.** In the $\mathfrak{Im}\,\mathrm{d}\neq 0$ case, or the case where $\mathrm{d}=0$ and $T<1$, we could also use the scheme described in Section 3.4. We prefer illustrate the method of contraction, where we will see that the counterpart of the singular contribution is an affine part. To simplify the exposition, we assume the problem one-dimensional (consider only the first component) and $\mathrm{d}>0$.

Since

$$\begin{aligned}\frac{\partial\mathcal{L}_\mathrm{d}}{\partial s}(T,s,0) &= \frac{\partial\mathcal{L}_0}{\partial s}(T,s+2\mathrm{d},0)\\ &= \mathcal{J}(1+s+2\mathrm{d}) - \mathcal{J}(1-T+s+2\mathrm{d})\\ &\quad -\mathcal{J}(1+\frac{s}{2}+\mathrm{d}) + \mathcal{J}(1-T+\frac{s}{2}+\mathrm{d})\end{aligned}$$

and since $\mathcal{J}(a+s)-\mathcal{J}(b+s) = (a-b)\log s + o(1)$ as $s\to\infty$, we have

$$\lim_{s\to\infty}\frac{\partial\mathcal{L}_\mathrm{d}}{\partial s}(s,0) = T\log 2$$

which corresponds to the endpoint of the interval allowed for $\mathfrak{Re}\log\Phi_n(T)$. On the other side, if $\mathrm{d}>0$,

$$\lim_{s\downarrow-(1-T)-2\mathrm{d}}\frac{\partial\mathcal{L}_d}{\partial s}(T,s,0) = \mathcal{J}(T) - 1 - \mathcal{J}\left(\frac{1+T}{2}\right) + \mathcal{J}\left(\frac{1-T}{2}\right) =: \xi_T.$$

There is a problem of non-steepness since it is not $-\infty$.

The first coordinate satisfies the LDP with good rate function

$$\xi\mapsto\inf\{h_0(T,\xi,\eta)|\eta\}$$

and from (3.4)

$$\inf\{h_0(T,\xi,\eta)|\eta\} = \inf\{\mathcal{H}_0(T,\varphi,\psi)|\varphi(T)=\xi\}$$

From the structure of $\mathcal{H}_0$ and H_a, it is clear that this infimum is achieved for $\dot\psi(.) = 0$ i.e. $\rho = 0$.

In the case of admissible φ, it remains to compute γ. Let us study the mapping $\gamma \mapsto \varphi(T,\gamma)$ where $\varphi(\cdot,\gamma)$ is given by (3.24) with $\rho = 0$. We follow the lines of argument of [22] pp. 3216-3218. We have

$$\frac{\partial\varphi(T,\gamma)}{\partial\gamma} = \log\frac{1+\gamma}{1-T+\gamma} - \frac{1}{2}\log\frac{1+\frac{\gamma}{2}}{1-T+\frac{\gamma}{2}} > 0\,. \tag{3.27}$$

The mapping $\gamma \mapsto \varphi(T,\gamma)$ is then bijective from $[-(1-T),\infty)$ to $[\xi_T, T\log 2)$.

Fixing $\xi \in (-\infty, T\log 2]$, let us look for optimal φ. Let $\gamma > -(1-T)$ (playing the role of a Lagrange multiplier). By the duality property, we have the inequality

$$(1-\tau)L^\star\left(\dot\varphi_a(\tau)\right) \geq \gamma\dot\varphi_a(\tau) - (1-\tau)L\left(\frac{\gamma}{1-\tau},0\right).$$

Using (3.2) and (3.22) we get, by integration of the above inequality:

$$\begin{aligned}
\mathcal{H}_0(T,\varphi,0) &= \int_0^T (1-\tau)L^\star(\dot\varphi_a(\tau))d\tau + \int_0^T (1-\tau)d\left(-\dot\varphi_s\right)(\tau) \\
&\geq \gamma\varphi_a(T) - \int_0^T (1-\tau)L\left(\frac{\gamma}{1-\tau},0\right)d\tau \\
&+ \int_0^T (1-\tau)d\left(-\dot\varphi_s\right)(\tau).
\end{aligned}$$

For every φ such that $\varphi(T) = \xi$, we then have

$$\begin{aligned}
\mathcal{H}_0(T,\varphi,0) &\geq \gamma\xi - \int_0^T (1-\tau)L\left(\frac{\gamma}{1-\tau},0\right)d\tau - \int_0^T (1-\tau+\gamma)d\dot\varphi_s(\tau) \\
&\geq \gamma\xi - \int_0^T (1-\tau)L\left(\frac{\gamma}{1-\tau},0\right)d\tau. \qquad (3.28)
\end{aligned}$$

We can now distinguish three cases:

• If $\xi \in [\xi_T, T\log 2)$, we choose the path v^ξ absolutely continuous and such that

$$\dot{v^\xi}(\tau) = \log(1-\tau+\gamma^\xi) - \log\left(1-\tau+\frac{\gamma^\xi}{2}\right),$$

where γ^ξ is uniquely determined by the condition $v^\xi(T) = \xi$. This path saturates the infimum and the expression of the action integral is clear.

• If $\xi < \xi_T$, set $\varepsilon = \xi_T - \xi$. Plugging $\gamma = -(1-T)$ in (3.28) yields for v such that $v(T) = \xi$:

$$\mathcal{H}_0(T,v,0) \geq -(1-T)\xi - \int_0^T (1-\tau)L\left(\frac{-(1-T)}{1-\tau},0\right) = (1-T)\varepsilon + h_0(T,\xi_T,0)$$

and this lower bound is achieved for the measure $\widetilde{v} = v^{\xi_T}(\tau)d\tau - \varepsilon\delta_T$, since

$$\mathcal{H}_a(v^{\xi_T}) = h_0(T,\xi_T,0)\ ,\ \int_0^T (1-\tau)\varepsilon\, d\delta_T(\tau) = (1-T)\varepsilon\,.$$

• If $\xi = T\log 2$, make $\xi = T\log 2$ in (3.28). We get, for all $\gamma > -(1-T)$,

$$h_0(T, T\log 2, 0) \geq \gamma T \log 2 - \int_0^T (1-\tau) L\left(\frac{\gamma}{1-\tau}, 0\right) d\tau. \tag{3.29}$$

When $\gamma \to \infty$, the integral is $F(1+\gamma) - F(1-T+\gamma) + F(1) - F(1-T) - 2F(1+\gamma/2) + 2F(1-T+\gamma/2)$ which tends to $-\infty$ as $-T\log\gamma$, so that finally, the RHS (3.29) tends to ∞ and we conclude $h_0(T, T\log 2, 0) = \infty$.

4. Connection with the spectral method

It can be interesting to connect the results of the previous section to the results obtained by looking directly at the empirical spectral distribution of the ensembles which are considered. This point of view is also discussed in [12].

The LDP for the empirical spectral distribution of the unitary ensemble is given in [11]. The rate function is the Voiculescu's logarithmic entropy:

$$I(\mu) = -\Sigma_{\mathbb{T}}(\mu) := -\iint_{\mathbb{T}\times\mathbb{T}} \log|z - z'|\ d\mu(z) d\mu(z'). \tag{4.1}$$

The circular Jacobi unitary ensemble yields also a LDP given in [5]. The rate function is

$$I_{\mathrm{d}}(\mu) = -\Sigma_{\mathbb{T}}(\mu) + \int_{\mathbb{T}} Q_{\mathrm{d}}(z) d\mu(z) + B(\mathrm{d})\,, \tag{4.2}$$

where

$$Q_{\mathrm{d}}(e^{\mathrm{i}\theta}) := -2(\mathfrak{Re}\,\mathrm{d})\log\left(2\sin\frac{\theta}{2}\right) - (\mathfrak{Im}\,\mathrm{d})(\theta - \pi)\ \ (\theta \in (0, 2\pi))\,. \tag{4.3}$$

and

$$\begin{aligned} B(\mathrm{d}) &= \int_0^1 \left[(x + 2\mathfrak{Re}\,\mathrm{d})\log(x + 2\mathfrak{Re}\,\mathrm{d}) - 2\mathfrak{Re}\left[(x+\mathrm{d})\log(x+\mathrm{d})\right]\right] dx \\ &+ \int_0^1 x \log \mathrm{d}x\,. \end{aligned} \tag{4.4}$$

(Notice that there was a mistake in [5], fixed in the arXiv version). With our notation, it yields:

$$B(\mathrm{d}) = F(1 + 2\mathfrak{Re}\,\mathrm{d}) - F(2\mathfrak{Re}\,\mathrm{d}) - 2\mathfrak{Re}\,F(1+\mathrm{d}) + 2\mathfrak{Re}\,F(\mathrm{d}) + F(1)\,.$$

If the mapping $\mu \in \mathcal{M}_1(\mathbb{T}) \mapsto \int \log(1-z)\, d\mu(z)$ were continuous, we would have by contraction:

$$h_{\mathrm{d}}(1, \xi, \eta) = \inf\{I_{\mathrm{d}}(\mu) \mid \mu \in \mathcal{M}_1(\mathbb{T}) : \int_{\mathbb{T}} \log(1-z)\, d\mu(z) = \xi + \mathrm{i}\eta\}\,. \tag{4.5}$$

We conjecture that this formula holds nevertheless, and we prove it in the one-dimensional case.

Proposition 4.1. *With the notation of Section 3, we have, for* $\mathrm{d} > 0$ *and for every* $\xi \in [0, \log 2)$

$$h_{\mathrm{d}}(1, \xi, 0) = \inf\{I_{\mathrm{d}}(\mu) \mid \int_{\mathbb{T}} \log|1-z|\, d\mu(z) = \xi\}\,. \tag{4.6}$$

We use the following interesting result.

Proposition 4.2. *We have*

$$\inf\{-\Sigma_{\mathbb{T}}(\mu) \mid \mu \in \mathcal{M}_1(\mathbb{T}) : \int_{\mathbb{T}} \log|1-z|\, d\mu(z) = \xi\} = -\Sigma_{\mathbb{T}}(\mu_a) \tag{4.7}$$

where

- *the measure $\mu_a \in \mathcal{M}_1(\mathbb{T})$ is defined by*

$$d\mu_a(z) = (1+a)\frac{\sqrt{\sin^2(\theta/2) - \sin^2(\theta_a/2)}}{2\pi\, \sin(\theta/2)} \mathbb{1}_{(\theta_a, 2\pi-\theta_a)}(\theta)\, d\theta \tag{4.8}$$

  where $z = e^{\mathrm{i}\theta}, \theta \in [0, 2\pi]$ and $\theta_a \in (0,\pi)$ is such that $\sin\theta_a/2 = \frac{a}{1+a}$.
- *a is the unique solution of*

$$\int_{\mathbb{T}} \log|1-z|\, d\mu_a(z) = \xi.$$

Moreover,

$$\int_{\mathbb{T}} \arg(1-z)\, d\mu_a(z) = 0\,.$$

The two propositions mean that to get the non standard mean value ξ for the logarithm of the determinant, the most probable way is to force the random operator to get an empirical spectral distribution close to μ_a.

Proof of Proposition 4.1 given Proposition 4.2. Let us first suppose that the Proposition 4.1 is true for $\mathrm{d} = 0$. One deduces, from this assumption and the definition (4.2) of I_{d}:

$$\begin{aligned}
&\inf\{I_{\mathrm{d}}(\mu) \mid \mu \in \mathcal{M}_1(\mathbb{T}) : \int \log|1-z| d\mu(z) = \xi\} \\
&= \inf\{-\Sigma_{\mathbb{T}}(\mu) \mid \mu \in \mathcal{M}_1(\mathbb{T}) : \int \log|1-z| d\mu(z) = \xi\} + B(\mathrm{d}) - 2d\xi \\
&= h_0(1,\xi,0) + B(\mathrm{d}) - 2\mathrm{d}\xi = h_{\mathrm{d}}(1,\xi,0),
\end{aligned}$$

the last equality coming from (3.5) and the fact that $B(\mathrm{d}) = -C_{\mathrm{d}}(1)$. Hence, it is sufficient to prove Proposition 4.1 for $\mathrm{d} = 0$.

The RHS of (4.6) is $-\Sigma_{\mathbb{T}}(\mu_a)$ and it remains to prove that it fits with the value

$$h_0(1,\xi,0) = \gamma\xi - F(1+\gamma) + F(\gamma) - F(1) + 2F\left(1+\frac{\gamma}{2}\right) - 2F\left(\frac{\gamma}{2}\right)$$

where in view of Theorem 3.4 (2), γ is the solution of (3.6) i.e.

$$\mathcal{J}(1+\gamma) - \mathcal{J}(\gamma) - \mathcal{J}\left(1+\frac{\gamma}{2}\right) + \mathcal{J}\left(\frac{\gamma}{2}\right) = \xi\,. \tag{4.9}$$

Applying [5] formula (5.25) in the special case $\mathrm{d} = a$ (real and positive), we see that, since the corresponding rate function vanishes for the limiting empirical measure μ_a,

$$\Sigma_{\mathbb{T}}(\mu_a) = \int_{\mathbb{T}} Q_a(\zeta)\, d\mu_a(\zeta) + B(a)\,, \tag{4.10}$$

where $Q_a(\zeta) = -2a\log|1-\zeta|$. Let us now compute:

$$I := \frac{2\pi}{1+a}\int_{\mathbb{T}} \log|1-\zeta|\; d\mu_a(\zeta)\,. \tag{4.11}$$

We have, by definition of μ_a in (4.8),

$$
\begin{aligned}
I &= \int_{\theta_a}^{2\pi-\theta_a} \log(2\sin(\theta/2)) \frac{\sqrt{\sin^2(\theta/2)-\sin^2(\theta_a/2)}}{\sin(\theta/2)}\, d\theta \\
&= 2\int_{\theta_a}^{\pi} \log(2\sin(\theta/2)) \frac{\sqrt{\sin^2(\theta/2)-\sin^2(\theta_a/2)}}{\sin(\theta/2)}\, d\theta\,, \qquad (4.12)
\end{aligned}
$$

for $\theta_a \in [0,\pi]$ and $\sin(\theta_a/2) = a/(1+a)$. The first change of variables $u = \sin(\theta/2)$ gives

$$
d\theta = \frac{2du}{\cos(\theta/2)} = \frac{2du}{\sqrt{1-u^2}},
$$

and then

$$
I = 4\int_{a/(1+a)}^{1} \log(2u) \sqrt{\frac{u^2-[a/(1+a)]^2}{1-u^2}} \frac{du}{u}.
$$

The second change of variable

$$
v = \sqrt{\frac{u^2-[a/(1+a)]^2}{1-u^2}} \Longleftrightarrow u = \sqrt{\frac{v^2+[a/(1+a)]^2}{1+v^2}},
$$

gives

$$
\begin{aligned}
\log(2u) &= \log 2 + \frac{1}{2}\log\left(v^2+[a/(1+a)]^2\right) - \frac{1}{2}\log(v^2+1), \\
\frac{du}{u} &= \left[\frac{v}{v^2+[a/(1+a)]^2} - \frac{v}{v^2+1}\right] dv.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
I &= 4\int_0^\infty \left[\log 2 + \frac{1}{2}\log\left(v^2+[a/(1+a)]^2\right) - \frac{1}{2}\log(v^2+1)\right]\cdot\left[\frac{v^2}{v^2+[a/(1+a)]^2} - \frac{v^2}{v^2+1}\right] dv \\
&= 2\int_{-\infty}^\infty \left[\log 2 + \frac{1}{2}\log\left(v^2+[a/(1+a)]^2\right) - \frac{1}{2}\log(v^2+1)\right]\cdot\left[\frac{1}{v^2+1} - \frac{[a/(1+a)]^2}{v^2+[a/(1+a)]^2}\right] dv. \qquad (4.13)
\end{aligned}
$$

Now, for $\alpha, \beta > 0$,

$$
\int_{-\infty}^\infty \frac{\alpha^2}{v^2+\alpha^2} dv = \alpha\pi, \qquad (4.14)
$$

$$
\int_{-\infty}^\infty \frac{\alpha^2\,\log(v^2+\beta^2)}{v^2+\alpha^2} dv = 2\alpha\pi\log(\alpha+\beta). \qquad (4.15)
$$

The first equality (4.14) is elementary. To get (4.15), one can differentiate the LHS with respect to β, and one obtains, for $\alpha \neq \beta$,

$$
\int_{-\infty}^\infty \frac{2\alpha^2\,\beta}{(v^2+\alpha^2)(v^2+\beta^2)} dv = \frac{2\alpha^2\,\beta}{\beta^2-\alpha^2}\left[\int_{-\infty}^\infty \frac{dv}{v^2+\alpha^2} - \int_{-\infty}^\infty \frac{dv}{v^2+\beta^2}\right] = \frac{2\pi\alpha}{\alpha+\beta}.
$$

By continuity, the last equality remains true for $\alpha = \beta$, and reversing the differentiation gives

$$
\int_{-\infty}^\infty \frac{\alpha^2\,\log(v^2+\beta^2)}{v^2+\alpha^2} dv = 2\alpha\pi\log(\alpha+\beta) + C(\alpha).
$$

Now, to determine $C(\alpha)$ let us write

$$\int_{-\infty}^{\infty} \frac{\alpha^2 \log(v^2+\beta^2)}{v^2+\alpha^2} dv = \alpha \log(\beta^2) \int_{-\infty}^{\infty} \frac{\alpha}{v^2+\alpha^2} dv + \alpha \int_{-\infty}^{\infty} \frac{\alpha \log(1+v^2/\beta^2)}{v^2+\alpha^2} dv\,.$$

The first integral is equal to $2\alpha\pi\log(\beta)$ and the second tends to zero for α fixed and $\beta \to \infty$ by dominated convergence. This gives $C(\alpha) = 0$ and then (4.15). Expanding (4.13), then using (4.14) and (4.15) gives:

$$I = \frac{2\pi}{1+a}\left[(1+2a)\log(1+2a) - (1+a)\log(1+a) - 2a\log(2a) + a\log(a)\right].$$

Coming back to the definition of I, we obtain:

$$\int_{\mathbb{T}} \log|1-\zeta|\; d\mu_a(\zeta) = \mathcal{I}(1+2a) - \mathcal{I}(1+a) - \mathcal{I}(2a) + \mathcal{I}(a) = \mathcal{E}_a(1)\,.$$

In particular,

$$\int_{\mathbb{T}} \log|1-\zeta|\; d\mu_{\gamma/2}(\zeta) = \xi,$$

and the value of a involved in Proposition 4.2 is equal to $\gamma/2$. It remains to check that $-\Sigma_{\mathbb{T}}(\mu_{\gamma/2}) = h_0(1,\xi,0)$. Indeed,

$$\begin{aligned} -\Sigma_{\mathbb{T}}(\mu_{\gamma/2}) &= -\int Q_{\gamma/2}(\zeta) d\mu_{\gamma/2}(\zeta) - B(\gamma/2) = \gamma\xi - B(\gamma/2) \\ &= \gamma\xi - F(1+\gamma) + F(\gamma) + 2F(1+\gamma/2) - 2F(\gamma/2) - F(1)\,, \end{aligned}$$

which proves Proposition 4.1. ∎

Remark 4.3. *The determination of I could also be viewed as a consequence of the weak convergence of the empirical spectral measure under $CJ_{na\beta'}$. Of course the mapping $\mu \mapsto \int_{\mathbb{T}} \log|1-\zeta|\; d\mu(\zeta)$ is not continuous, but since the extremal eigenvalues converge to the extreme points of the support of μ_a (see* [2]*), we have*

$$\int_{\mathbb{T}} \log|1-\zeta|\, d\mu_a(\zeta) = \lim_{n\to\infty} \frac{1}{n} \log|\Phi_n(1)| = \mathcal{E}_a(1)$$

where the last equality comes from (2.9) in Theorem 2.1.

Proof of Proposition 4.2 A possible way consists in studying the maximization of the functional

$$\Sigma_{\mathbb{T}}(\mu) + r\int_{\mathbb{T}} \log|1-z|\; d\mu(z)\,,$$

where r is some Lagrange multiplier.

It is convenient to push-forward this problem to $\mathbb{R}$ via the Cayley transformation

$$z = e^{\mathrm{i}\theta} = \frac{\lambda+\mathrm{i}}{\lambda-\mathrm{i}}$$

so that, for $z, z' \neq 1$,

$$\begin{aligned} \log|1-z| &= -\frac{1}{2}\log(1+\lambda^2) + \log 2 \\ \log|z-z'| &= \log|\lambda-\lambda'| - \frac{1}{2}\log(1+\lambda^2) - \frac{1}{2}\log(1+\lambda'^2) + \log 2\,. \end{aligned}$$

Hence, one needs to minimize:

$$\Sigma_{\mathbb{R}}(\nu) + 2\int_{\mathbb{R}} Q(\lambda)\, d\nu(\lambda),$$

over $\nu \in \mathcal{M}_1(\mathbb{R})$, where

$$\Sigma_{\mathbb{R}}(\nu) = \iint_{\mathbb{R}\times\mathbb{R}} \log\left(\frac{1}{|\lambda-\lambda'|}\right) d\nu(\lambda) d\nu(\lambda')$$

and

$$2Q(x) = \big(1+\frac{r}{2}\big)\log(1+x^2).$$

This problem can be connected to the research of the equilibrium measure in the Cauchy ensemble, i.e. for the Coulomb gas with probability distribution given by formula (3.124) in Forrester [10]. According to Saff and Totik [24], this potential Q is admissible (p.27) as soon as $r > 0$. In that case the minimizer is unique (Theorem 1.3), its support is compact. In the book, an explicit method is presented for a class of potentials satisfying some conditions, in particular if Q is even, differentiable and such that $xQ'(x)$ is positive and increasing in $(0,\infty)$. These properties are satisfied here since:

$$xQ'(x) = \big(1+\frac{r}{2}\big)\frac{x^2}{1+x^2}\,.$$

Then ([24] Corollary 1.12 p. 203) the support is $S = [-b, b]$ where b is solution of

$$\frac{2}{\pi}\int_0^1 \frac{btQ'(bt)}{\sqrt{1-t^2}}\, dt = 1 \tag{4.16}$$

i.e.

$$\frac{2+r}{\pi}\int_0^1 \frac{b^2t^2}{(1+b^2t^2)\sqrt{1-t^2}}\, dt = 1$$

or

$$\int_0^1 \frac{dt}{(1+b^2t^2)\sqrt{1-t^2}} = \frac{\pi r}{2(2+r)} \tag{4.17}$$

which gives (make $t = (1+\tau^2)^{-1/2}$)

$$\int_0^\infty \frac{d\tau}{1+b^2+\tau^2} = \frac{\pi r}{2(2+r)}$$

i.e.

$$b = \frac{2\sqrt{1+r}}{r}\,.$$

To find the extremal measure, we apply Theorem IV.3.1 in [24], which contains the following result:

Theorem 4.4 (Lubinsky-Saff)**.** *Let f be a differentiable even function on $[-1,+1]$, such that $sf'(s)$ is increasing for $s \in (0,1)$ and for some $1 < p < 2$, $f'(s)/\sqrt{1-s^2} \in L^p[-1,1]$. Then the integral equation*

$$\int_{-1}^1 \log\frac{1}{|x-t|} g(t) dt = -f(x) + C_f, \quad x \in (-1,1)\,,$$

(where C_f is some constant) has a solution of the form

$$g(t) = L[f'](t) + \frac{B_f}{\pi\sqrt{1-t^2}} \tag{4.18}$$

where

$$L[f'](t) = \frac{2}{\pi^2}\int_0^1 \frac{\sqrt{1-t^2}\,(sf'(s)-tf'(t))}{\sqrt{1-s^2}\,(s^2-t^2)}\,ds, \tag{4.19}$$

$$B_f = 1 - \frac{1}{\pi}\int_{-1}^1 \frac{sf'(s)}{\sqrt{1-s^2}}\,ds\,.$$

Here, we look for the measure μ with support $[-b,b]$ such that for $z \in (-b,b)$,

$$\int_{-b}^b \log\frac{1}{|z-t|}d\mu(t) = -Q(z) + C. \tag{4.20}$$

Theorem 4.4 is set for $b = 1$, so we need a scaling. Equation (4.20) becomes, with $d\mu(x) = g_b(x)dx$,

$$\int_{-1}^1 \log\frac{1}{|z-t|} bg_b(bt)dt = -Q(bz) + C'. \tag{4.21}$$

Theorem 4.4 can now be applied by taking $f(s) = Q(bs)$, and one has $bg_b(bt) = g(t)$, $B_f = 0$ and

$$\frac{sf'(s)-tf'(t)}{s^2-t^2} = b^2\Big(1+\frac{r}{2}\Big)\frac{1}{(1+b^2t^2)(1+b^2s^2)}\,.$$

Owing to (4.17), we get

$$bg_b(bt) = \frac{\left(1+\sqrt{1+b^2}\right)}{\pi}\,\frac{\sqrt{1-t^2}}{(1+b^2t^2)}\,\mathbb{I}_{[-1,1]}(t)$$

i.e.

$$g_b(x) = \frac{\left(1+\sqrt{1+b^2}\right)}{b\pi}\,\frac{\sqrt{1-x^2b^{-2}}}{(1+x^2)}\,\mathbb{I}_{[-b,b]}(x).$$

Now, it is enough to apply Theorem I.3.3 of [24] as follows. The support of the equilibrium measure is $S = [-b,b]$, the measure μ is supported by S, has a finite logarithmic energy and $\int_{-b}^b \log\frac{1}{|z-t|}d\mu(t) + Q(z)$ is constant for $z \in S$, so μ is the equilibrium measure. To find μ_a, it is enough to carry μ back on the circle by the inverse transformation

$$\lambda = \frac{1}{i}\frac{1+z}{1-z}\,.$$

∎

5. Appendix

5.1. Some properties of $\ell = \log\Gamma$ and $\Psi = \ell'$. From the Binet formula (Abramowitz and Stegun [1] or Erdélyi et al. [9] p.21), we have for $\mathfrak{Re}\, x > 0$

$$\ell(x) = (x-\frac{1}{2})\log x - x + 1 + \int_0^\infty f(s)[e^{-sx}-e^{-s}]\,ds \tag{5.1}$$

$$= (x-\frac{1}{2})\log x - x + \frac{1}{2}\log(2\pi) + \int_0^\infty f(s)e^{-sx}\,ds\,, \tag{5.2}$$

where the function f is defined by

$$f(s) = \left[\frac{1}{2}-\frac{1}{s}+\frac{1}{e^s-1}\right]\frac{1}{s} = 2\sum_{k=1}^\infty \frac{1}{s^2+4\pi^2k^2}\,, \tag{5.3}$$

and satisfies for every $s \geq 0$:

$$0 < f(s) \leq f(0) = 1/12\ ,\quad 0 < \left(sf(s) + \frac{1}{2}\right) < 1\,. \tag{5.4}$$

By differentiation (recall that Ψ is the Digamma function Γ'/Γ),

$$\log x - \Psi(x) = \frac{1}{2x} + \int_0^\infty sf(s)e^{-sx}\,ds = \int_0^\infty e^{-sx}\left(sf(s) + \frac{1}{2}\right)ds\,. \tag{5.5}$$

Moreover, since $\Psi(x+1) = \frac{1}{x} + \Psi(x)$, we have a variation of (5.5):

$$\begin{aligned} \log x - \Psi(x+1) &= -\frac{1}{2x} + \int_0^\infty sf(s)e^{-sx}\,ds \\ &= \int_0^\infty e^{-sx}\left(sf(s) - \frac{1}{2}\right)ds\,. \end{aligned} \tag{5.6}$$

As easy consequences, we have, for every $x > 0$,

$$0 \quad < \quad x\left(\log x - \Psi(x)\right) \leq 1\,, \tag{5.7}$$

$$0 \quad < \quad \log x - \Psi(x) - \frac{1}{2x} \leq \frac{1}{12x^2}\,. \tag{5.8}$$

Differentiating again we see that for $q \geq 1$, $\mathfrak{Re}\, x > 0$,

$$\Psi^{(q)}(x) = (-1)^{q-1}(q-1)!x^{-q} + (-1)^{q-1}\int_0^\infty e^{-sx}s^q\left(sf(s) + \frac{1}{2}\right)\,ds$$

and then

$$|\Psi^{(q)}(x) - (-1)^{q-1}(q-1)!x^{-q}| \leq (\mathfrak{Re}\, x)^{-q-1}q!\,. \tag{5.9}$$

Another useful formula is

$$\Psi(z+1) = -\gamma - \sum_{k=1}^{\infty}\left(\frac{1}{k+z} - \frac{1}{k}\right)\,, \tag{5.10}$$

fir $z + 1 \notin \mathbb{R}_-$.

5.2. **The density $g_r^{(\delta)}$ and some moments related to it.** Recall that for $r > 0$ and δ such that $r + 2\mathfrak{Re}\,\delta + 1 > 0$, the density $g_r^{(\delta)}$ on the unit disc $\mathbb{D}$ is given by

$$g_r^{(\delta)}(z) = c_{r,\delta}(1 - |z|^2)^{r-1}(1-z)^{\bar{\delta}}(1-\bar{z})^{\delta}$$

where $c_{r,\delta}$ is the normalization constant. The following lemma is the key to compute $c_{r,\delta}$ and the moments of $g_r^{(\delta)}$.

Lemma 5.1. *Let s, t, ℓ be complex numbers such that: $\mathfrak{Re}\,\ell$, $\mathfrak{Re}(s+\ell+1)$, $\mathfrak{Re}(t+\ell+1)$ and $\mathfrak{Re}(s+t+\ell+1)$ are strictly positive. Then, the following identity holds:*

$$\int_{\mathbb{D}}(1-|z|^2)^{\ell-1}(1-z)^s(1-\bar{z})^t\mathrm{d}^2z = \pi\Gamma\begin{bmatrix}\ell\ ,\ \ell+1+s+t\\ \ell+1+s\ ,\ \ell+1+t\end{bmatrix}\,, \tag{5.11}$$

where for the sake of simplicity we use the polygamma symbol

$$\Gamma\begin{bmatrix}a, b, \cdots\\ c, d, \cdots\end{bmatrix} := \frac{\Gamma(a)\Gamma(b)\cdots}{\Gamma(c)\Gamma(d)\cdots}.$$

A proof of this result is given in [5]. A first consequence is that

$$c_{r,\delta} = \pi^{-1}\Gamma\begin{bmatrix} r+1+\delta \ , \ r+1+\bar{\delta} \\ r \ , \ r+1+\delta+\bar{\delta}\end{bmatrix} . \tag{5.12}$$

A second consequence is that if γ has the density $g_r^{(\delta)}$, then we have

$$\mathbb{E}(1-\gamma)^a(1-\bar{\gamma})^b = \Gamma\begin{bmatrix} r+1+\delta+\bar{\delta}+a+b \ , \ r+1+\bar{\delta} \ , \ r+1+\delta \\ r+1+\delta+\bar{\delta} \ , \ r+1+\bar{\delta}+a \ , \ r+1+\delta+b\end{bmatrix} \tag{5.13}$$

as soon as all the real parts of the arguments of the gamma functions are strictly positive.

Let us notice that for $r=0$ the RHS of (5.13) is the Mellin-Fourier transform of $1-\gamma$ when $\gamma \in \mathbb{T}$ is distributed according to $\lambda^{(\delta)}$.

In this paper, we need the following computations, in order to deduce the moments of $\log(1-\gamma)$. The quantities involved below are all well-defined as soon as $s > s_0$, where s_0 is some strictly negative quantity depending on r and δ, and in particular, for (s,t) in the neighborhood of $(0,0)$, one can write:

$$\begin{aligned}\Lambda(s,t) \quad &:= \quad \log\mathbb{E}\exp\left(2s\mathfrak{Re}\log(1-\gamma)+2t\mathfrak{Im}\log(1-\gamma)\right) = \\ &= \quad \log\mathbb{E}\exp\left(\mathfrak{Re}\left(2(s-\mathrm{i}t)\log(1-\gamma)\right)\right) = \\ &= \quad \log\mathbb{E}(1-\gamma)^{s-\mathrm{i}t}(1-\bar{\gamma})^{s+\mathrm{i}t} = \\ &= \quad \ell\big(r+1+\delta+\overline{\delta}+2s\big)-\ell\big(r+1+\delta+\overline{\delta}\big) \\ &\qquad -\ell\left(r+1+\overline{\delta}+s-\mathrm{i}t\right)-\ell\Big(r+1+\delta+s+\mathrm{i}t\Big) \\ &\qquad +\ell\big(r+1+\overline{\delta}\big)+\ell\big(r+1+\delta\big).\end{aligned} \tag{5.14}$$

To compute moments we need differentiation. First we have:

$$\begin{aligned}\frac{\partial}{\partial s}\Lambda(s,t) \quad &= \quad 2\Psi\left(r+1+\delta+\overline{\delta}+2s\right) \\ &\qquad -\Psi\left(r+1+\delta+s+\mathrm{i}t\right)-\Psi\left(r+1+\overline{\delta}+s-\mathrm{i}t\right) \\ \frac{\partial}{\partial t}\Lambda(s,t) \quad &= \quad \mathrm{i}\Psi\Big(r+1+\bar{\delta}+s-\mathrm{i}t\Big)-\mathrm{i}\Psi\Big(r+1+\delta+s+\mathrm{i}t\Big).\end{aligned} \tag{5.15}$$

The first moment is then:

$$\begin{aligned}\mathbb{E}\,\mathfrak{Re}\log(1-\gamma) \quad &= \quad \Psi(r+1+\delta+\bar{\delta})-\frac{1}{2}\Psi(r+1+\delta)-\frac{1}{2}\Psi(r+1+\bar{\delta}) \\ \mathbb{E}\,\mathfrak{Im}\log(1-\gamma) \quad &= \quad \frac{1}{2\mathrm{i}}\Psi(r+1+\delta)-\frac{1}{2\mathrm{i}}\Psi(r+1+\bar{\delta})\end{aligned}$$

or

$$\mathbb{E}\log(1-\gamma) = \Psi(r+1+\delta+\bar{\delta})-\Psi(r+1+\bar{\delta})\,. \tag{5.16}$$

Differentiating again (5.15) we get

$$\begin{aligned}
\frac{\partial^2}{\partial s^2}\Lambda(s,t) &= 4\Psi'\big(r+1+\delta+\overline{\delta}+2s\big) \\
&\quad -\Psi'\Big(r+1+\delta+s+\mathrm{i}t\Big)-\Psi'\Big(r+1+\overline{\delta}+s-\mathrm{i}t\Big) \\
\frac{\partial^2}{\partial t^2}\Lambda(s,t) &= \Psi'\Big(r+1+\bar{\delta}+s-\mathrm{i}t\Big)+\Psi'\Big(r+1+\delta+s+\mathrm{i}t\Big) \\
\frac{\partial^2}{\partial s\partial t}\Lambda(s,t) &= -\mathrm{i}\Psi'\Big(r+1+\delta+s+\mathrm{i}t\Big)+\mathrm{i}\Psi'\Big(r+1+\bar{\delta}+s-\mathrm{i}t\Big)
\end{aligned} \tag{5.17}$$

and the second moments are

$$\begin{aligned}
\operatorname{Var}\mathfrak{Re}\log(1-\gamma) &= \Psi'\big(r+1+\delta+\overline{\delta}\big)-\frac{1}{4}\Psi'\big(r+1+\delta\big)-\frac{1}{4}\Psi'\big(r+1+\overline{\delta}\big) \\
\operatorname{Var}\mathfrak{Im}\log(1-\gamma) &= \frac{1}{4}\Psi'\big(r+1+\delta\big)+\frac{1}{4}\Psi'\big(r+1+\overline{\delta}\big) \\
\operatorname{Cov}\left(\mathfrak{Re}\log(1-\gamma),\mathfrak{Im}\log(1-\gamma)\right) &= \frac{1}{4\mathrm{i}}\Psi'\big(r+1+\delta\big)-\frac{1}{4\mathrm{i}}\Psi'\big(r+1+\bar{\delta}\big).
\end{aligned} \tag{5.18}$$

5.3. **Complex logarithm and characteristic polynomial.** Let E_k be the set of the complex $k\times k$ matrices with no eigenvalue on the interval $[1,\infty)$. For $V\in E_k$, let us define

$$\log\det(I_k-V):=\sum_{j=1}^{k}\log(1-\lambda_j),$$

where the λ_j's are the roots, counted with multiplicity, of the polynomial $z\mapsto\det(zI_k-V)$, and where in the right-hand side, one considers the principal branch of the logarithm. This definition is meaningful, since by assumption, $1-\lambda_j\notin\mathbb{R}_-$ for all $j\in\{1,\dots,k\}$. By the continuity of the set of roots of a polynomial with respect to its coefficients, the set E_k is open and the function $V\mapsto\log\det(I_k-V)$ defined just above is continuous on E_k. In fact, since E_k is connected (this is easily checked by tridiagonalizing the matrices), this is the unique way to define the logarithm of $\det(I_k-V)$ as a continuous function of $V\in E_k$ if we assume that it should take the value zero at $V=0$.

Now, with the notation of the beginning of the paper, the matrix $G_k(U_n)$ is a submatrix of the unitary matrix U_n, and all its eigenvalues have modulus bounded by 1. If we assume $\gamma_0,\dots,\gamma_{n-1}\neq 1$ (which holds almost surely under $\mathrm{CJ}^{(n)}_{\beta,\delta}$), then by (1.2), $\Phi_{k,n}(1)\neq 0$, and one easily deduces that $G_k(U_n)\in E_k$, which allows to define $\log\Phi_{k,n}(1)$ without ambiguity. Now, the map from $\mathbb{D}^{n-1}\times(\mathbb{U}\backslash\{1\})$ to $\mathbb{R}$, given by

$$(\gamma_1,\dots,\gamma_{n-1})\mapsto\sum_{j=0}^{k-1}\log(1-\gamma_j)$$

is continuous if we take the principal branch of the logarithm, and since U_n depends continuously on $(\gamma_1,\dots,\gamma_{n-1})\in\mathbb{D}^{n-1}\times(\mathbb{U}\backslash\{1\})$, as it can be checked in [5], the map

$$(\gamma_1,\dots,\gamma_{n-1})\mapsto\log\Phi_{k,n}(1)$$

is also continuous. These two maps have the same exponential, and one can check that they are both real if the γ_j's are all real. Hence, they are equal, which fully justifies the equation

$$\log \Phi_{k,n}(1) = \sum_{j=0}^{k-1} \log(1-\gamma_j). \tag{5.19}$$

5.4. Abel-Plana summation formula.

Theorem 5.2. *Let $m < n$ be integers and let g be a holomorphic function on the strip $\{t \in \mathbb{C}, n \leq \mathfrak{Re}\, t \leq m\}$ (i.e. g is continuous on this strip and holomorphic in its interior). We assume that $g(t) = o\left(\exp(2\pi|\mathfrak{Im}\, t|)\right)$ as $\mathfrak{Im}\, t \to \pm\infty$, uniformly with respect to $\mathfrak{Re}\, t \in [n, m]$. Then,*

$$\begin{aligned} \sum_{j=m+1}^{n} g(j) &= \int_m^n g(t)dt + \frac{g(n)-g(m)}{2} \\ &+ \mathrm{i}\int_0^\infty \frac{g(m+\mathrm{i}y)-g(n+\mathrm{i}y)-g(m-\mathrm{i}y)+g(n-\mathrm{i}y)}{e^{2\pi y}-1}dy\,. \end{aligned} \tag{5.20}$$

For a proof see [20] p. 290.

References

[1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York, 1972. pp. 258-259, 9th edition.

[2] G. Borot and A. Guionnet. Asymptotic expansion of β matrix models in the one-cut regime. *Comm. Math. Phys.*, 317(2):447–483, 2013.

[3] P. Bourgade, C. P. Hughes, A. Nikeghbali, and M. Yor. The characteristic polynomial of a random unitary matrix: a probabilistic approach. *Duke Math. J.*, 145(1):45–69, 2008.

[4] P. Bourgade, A. Nikeghbali, and A. Rouault. The characteristic polynomial on compact groups with Haar measure: some equalities in law. *C. R. Math. Acad. Sci. Paris*, 345(4):229–232, 2007.

[5] P. Bourgade, A. Nikeghbali, and A. Rouault. Circular Jacobi ensembles and deformed Verblunsky coefficients. *Int. Math. Res. Not. IMRN*, (23):4357–4394, 2009.

[6] P. Bourgade, A. Nikeghbali, and A. Rouault. Ewens Measures on Compact Groups and Hypergeometric Kernels. In *Séminaire de probabilités XLIII*, volume 2006 of *Lecture Notes in Math.*, pages 351–378. Springer, Berlin, 2010.

[7] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Springer, 2nd edition, 1998.

[8] H. Dette and F. Gamboa. Asymptotic properties of the algebraic moment range process. *Acta Math. Hungar.*, 116:247–264, 2007.

[9] A. Erdelyi, W. Magnus, F. Oberhettinger, and F.G. Tricomi. *Higher transcendental functions*, volume I. Krieger, New-York, 1981.

[10] P. J. Forrester. *Log-gases and random matrices*, volume 34 of *London Mathematical Society Monographs Series.* Princeton University Press, Princeton, NJ, 2010.

[11] F. Hiai and D. Petz. A large deviation theorem for the empirical eigenvalue distribution of random unitary matrices. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(1):71–85, 2000.

[12] C.P. Hughes, J.P. Keating, and N. O'Connell. On the Characteristic Polynomial of a Random Unitary Matrix. *Comm. Math. Phys.*, 220(2):429–451, 2001.

[13] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes.* Springer, 1987.

[14] J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2+it)$. *Comm. Math. Phys.*, 214(1):57–89, 2000.

[15] R. Killip. Gaussian fluctuations for β ensembles. *Int. Math. Res. Not.*, 2008(8):Art. ID rnn007, 19, 2008.

[16] R. Killip and I. Nenciu. Matrix models for circular ensembles. *Int. Math. Res. Not.*, 50:2665–2701, 2004.
[17] R. Killip and M. Stoiciu. Eigenvalue statistics for CMV matrices: from Poisson to clock via random matrix ensembles. *Duke Math. J.*, 146(3):361–399, 2009.
[18] Y. A. Neretin. Hua-type integrals over unitary groups and over projective limits of unitary groups. *Duke Math. J.*, 114(2):239–266, 2002.
[19] Yu. A. Neretin. Matrix analogues of the B-function, and the Plancherel formula for Berezin kernel representations. *Mat. Sb.*, 191(5):67–100, 2000.
[20] F.W.J. Olver. *Asymptotics and special functions*. Academic Press, New York, 1974.
[21] R.T. Rockafellar. Integrals which are convex functionals, II. *Pacific J. Math.*, 39(2):439–469, 1971.
[22] A. Rouault. Asymptotic behavior of random determinants in the Laguerre, Gram and Jacobi ensembles. *ALEA Lat. Am. J. Probab. Math. Stat.*, 3:181–230 (electronic), 2007.
[23] E. Ryckman. Linear statistics of point processes via orthogonal polynomials. *J. Stat. Phys.*, 132(3):473–486, 2008.
[24] E.B. Saff and V. Totik. *Logarithmic potentials with external fields*, volume 316 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997. Appendix B by Thomas Bloom.
[25] B. Simon. *Orthogonal polynomials on the unit circle. Part 1: Classical theory*. Colloquium Publications. American Mathematical Society 54, Part 1. Providence, RI: American Mathematical Society (AMS), 2005.
[26] B. Simon. CMV matrices: Five years after. *J. Comput. Appl. Math.*, 208:120–154, 2007.

Institut de Mathématiques de Toulouse, Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse Cedex 9, France
E-mail address: joseph.najnudel@math.univ-toulouse.fr

Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
E-mail address: ashkan.nikeghbali@math.uzh.ch

Université Versailles-Saint Quentin, LMV, Bâtiment Fermat, 45 avenue des Etats-Unis, 78035 Versailles Cedex
E-mail address: alain.rouault@uvsq.fr